

FOURTH EDITION

RNA

METHODOLOGIES



ROBERT FARRELL



RNA Methodologies

This page intentionally left blank

RNA Methodologies

A Laboratory Guide for Isolation and Characterization

4th Edition

Robert E. Farrell, Jr., Ph.D
The Pennsylvania State University
York, PA



ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK
OXFORD • PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE
SYDNEY • TOKYO

Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

Fourth edition, 2010

Copyright © 2010 Elsevier Inc. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively, visit the Science and Technology Books website at www.elsevierdirect.com/rights for further information

Notice No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-374727-3

For information on all Academic Press publications
visit our website at elsevierdirect.com

Typeset by Macmillan Publishing Solutions (www.macmillansolutions.com)

Printed and bound United States of America

10 11 12 13 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

*For Catherine Ann,
Sean Patrick, Emma Catherine, Liam Michael, and Patrick Joseph*

This page intentionally left blank

Contents

Preface	xxi
1 RNA and the Cellular Biochemistry Revisited	1
Why Study RNA?	1
What is RNA?	2
Assembly of Polynucleotides	5
Types of RNA	8
Transcription and the Central Dogma	10
Important Plant and Animal Models for Studying Transcription	11
Promoters and Regulatory Elements	12
Gene and Genome Organization Affect Transcription	14
RNA Polymerases and the Products of Transcription	18
Messenger RNA	22
Topology of a Typical mRNA Molecule	23
5' Cap	23
Leader Sequence	25
Coding Region	25
Trailer Sequence	26
Poly(A) Tail	26
Organellar mRNAs	28
mRNA Stability, Transport, and Turnover	29
Bicistronic mRNAs	30
Prokaryotic mRNAs	30
mRNA Sequence and Structure Affect Translation	31
Levels of Gene Regulation	32
Alternative Splicing of mRNA from a Single Genetic Locus	34
<i>Trans</i> -splicing; mRNA Repair	35
Overview of Small RNAs	36
miRNA Regulation of Gene Expression	37
References	38
2 RNA Isolation Strategies	45
Rationale	45
Goals in the Purification of RNA	47
Lysis Buffer Formulations	49
Gentle Lysis Buffers	49

Protocol: Isolation of Cytoplasmic RNA by Hypotonic Lysis	51
Chaotropic Lysis Buffers	56
Isolation of RNA with Guanidinium Buffers	57
Guanidinium–Acid–Phenol Extraction Techniques	58
Protocol: Guanidinium–Acid–Phenol Extraction	59
Density Gradient Centrifugation	60
Cesium Chloride	61
Protocol: Cesium Chloride (CsCl) Gradients	61
Cesium Trifluoroacetate (CsTFA)	65
Protocol: Cesium Trifluoroacetate (CsTFA) Gradients	65
Simultaneous Isolation of RNA and DNA	68
Protocol: Simultaneous Isolation of RNA and DNA	69
The Word on Kits	71
Silica Technology	72
Isolation of Cytoplasmic RNA on a Silica Column	73
Affinity Matrices	73
Other Methods	74
Protocol: Rapid Isolation of RNA with SDS and Potassium Acetate Reagents	74
Protocol: Isolation of Prokaryotic RNA	75
Protocol: Isolation of RNA from Yeast	76
Short- and Long-Term Storage of Purified RNA	78
References	79
3 The Truth about Tissues	81
Rationale	81
Tissue Culture or Tissue?	81
Advantages of Cell Culture	82
Advantages of Tissue Samples	83
Homogenization Methods	84
Polytron Disruption	84
Dounce Homogenization	86
BeadBeater™ Technique	87
RNA Isolation Strategies for Various Organs and Tissues	87
Fresh Tissue	89
Frozen Tissue	90
Fixed Tissue	91
Protocol: LiCl–Urea Method for RNA Isolation from Tissue	92
Protocol: RNA Isolation from Lipid-Enriched Samples	95
Purification of Polysome-Engaged mRNA	96
Protocol: Isolation of Polysomal mRNA	98
Collecting Samples in the Field	99
RNA “Clean-Up” Methods	100

Troubleshooting RNA Isolation from Tissue	101
References	102
4 Going Green: RNA and the Molecular Biology of Plants	105
Rationale	105
RNA Isolation and the Peculiarities of Plants	105
Types of RNA Produced in Plant Cells	109
Protocol: RNA Isolation from Leaf	110
Protocol: RNA Isolation from Bark	112
Protocol: RNA Isolation from Fruit	113
Strategies for RNA Isolation from Other Plant Tissues	115
Troubleshooting RNA Isolation from Plant Tissue	116
References and Suggested Reading	117
5 Isolation of Polyadenylated RNA	121
Rationale	121
Polyadenylation	122
The Poly(A) Caveat	123
Example 1	124
Example 2	124
Selection of Polyadenylated Molecules	126
Magnetic Bead Technology for Poly(A) ⁺ Purification	127
Oligo(dT)-Cellulose Column Chromatography	129
Protocol: Purification of Biophysical Quantities of Poly(A) ⁺ RNA	130
Rapid, Non-Column Poly(A) ⁺ Purification	135
Protocol: Non-Column Poly(A) ⁺ Purification	135
References	136
6 Quality Control for RNA Preparations	139
Rationale	139
Quality Control Technique 1: UV Spectrophotometry and Absorption Ratios	140
Spectrophotometric Methods	140
Determination of Nucleic Acid Concentration	141
Determination of Nucleic Acid Purity	143
Non-spectrophotometric Methods	145
Quality Control Technique 2: Electrophoretic Profile of the RNA Protocol	147
Quality Control Technique 3: UV Shadowing Protocol	150
Quality Control Technique 4: Sample Capacity to Support RT-PCR	152
Quality Control Technique 5: Northern Analysis	153
Quality Control Technique 6: Sample Capacity to Support <i>In Vitro</i> Translation	153
References	154

7 Resilient Ribonucleases	155
Rationale	155
Elimination of Ribonuclease Activity	156
Latent RNase Contamination Issues	157
Types of Ribonuclease Inhibitors	158
Specific Inhibitors	158
Vanadyl Ribonucleoside Complex (VDR; VRC)	158
RNasin®	159
Non-specific Inhibitors	160
Preparation of Equipment and Reagents	161
Diethyl Pyrocarbonate (DEPC)	162
Alternative: Sterile Water	164
Hydrogen Peroxide	164
NaOH and SDS	165
Other Reagents Used to Control Nuclease Activity	165
Guanidine Hydrochloride	165
Guanidine Thiocyanate	165
Sodium Dodecyl Sulfate	166
<i>N</i> -Laurylsarcosine	166
Phenol:Chloroform:Isoamyl Alcohol	166
8-Hydroxyquinoline	167
Cesium Chloride	167
Cesium Trifluoroacetate	168
Proteinase K	168
RNAlater®	169
Protocol: Synthesis of VDR	169
References	170
8 Stringency: Conditions that Influence Nucleic Acid Structure	173
Types of Double-Stranded Molecules	173
Importance of Controlling Stringency	174
Effect of Salt on Stringency	176
Effect of pH on Stringency	176
Effect of Temperature on Stringency	177
Effect of Formamide on Stringency	177
Effect of Urea on Stringency	177
References	178
9 Electrophoresis of RNA	179
Rationale	179
Normalization of Nucleic Acids	180
Protocol: Poly(A) Normalization	183
Sample Preparation	183
RNA Denaturing Systems for Agarose Gel Electrophoresis	185
Formaldehyde Denaturation	187

Protocol: Formaldehyde Denaturing Gels	188
Urea Denaturation	189
Protocol: Urea Denaturation	191
Glyoxal/Dimethyl Sulfoxide Denaturation	191
Protocol: Glyoxalation and Electrophoresis of RNA	192
Running RNA on Non-denaturing Gels	194
Molecular Weight Standards	194
Proper Use of Size Standards	195
Ribosomal RNA	197
Gel Staining Techniques	204
Ethidium Bromide	204
SYBR [®] Green	207
SYBR [®] Gold	207
SYBR [®] Safe	208
GelStar [®]	209
Silver Staining	209
Acridine Orange	210
Methylene Blue	210
Safety Considerations and Equipment Maintenance	212
Running Agarose Gels for the First Time: A Few Tips	213
Essential Vocabulary	213
Points to Keep in Mind	213
References	216
10 Photodocumentation and Image Analysis	221
Rationale	221
Safety First	222
Digital Image Analysis	223
Image Formats	228
Practical Considerations	229
Digital Image Analysis for Every Budget	231
Image Analysis Workshops	232
PhosphorImagers	232
Traditional Methods of Photodocumentation	233
Sample Visualization	234
Filtration	234
Tips for Optimizing Electrophoretograms	235
Inherent Limitations of Photographic and X-ray Films	238
References and Suggested Reading	239
11 Northern Analysis	241
Rationale	241
Choice of Filter Membrane	242
Nitrocellulose	243
Nylon	244
Polyvinylidene Difluoride	245

Handling and Filter Preparation	245
Northern Transfer Techniques	246
Capillary Transfer	246
TurboBlotter™	247
Vacuum Blotting	248
Electroblotting	249
Alkaline Blotting	249
Protocol: RNA Transfer by Passive Capillary Diffusion	250
Protocol: TurboBlotter Downward Transfer of RNA	252
Post-Transfer Handling of Filters	254
Formaldehyde Denaturing Systems	254
Glyoxal Denaturing Systems	254
Option 1	254
Option 2	255
Immobilization Techniques	255
Baking	255
Crosslinking by UV Irradiation	256
Protocol: UV Crosslinking RNA to Nylon Filters	257
Post-Fixation Handling of Filters	257
Reverse Northern Analysis	258
References	259
12 Nucleic Acid Probe Technology	261
Rationale	261
Probe Classification	263
Selection of Labeling System	263
Isotope Labeling	265
Minimizing Decomposition Problems	266
Non-Isotopic Labeling	267
The Ubiquitous Dyes Cy3 and Cy5	267
Popular Chemiluminescence Formats	268
Biotin	268
Digoxigenin	270
Fluorescein	271
Direct Enzyme Labeling	271
DNA Probes	272
DNA Probe Synthesis	273
Polymerase Chain Reaction	274
Random Priming	274
Nick Translation	274
5' End-Labeling	275
3' End-Labeling	275
Antisense RNA Probes	276
Characteristics of RNA Probes	278
RNA Probe Synthesis	278
<i>In Vitro</i> transcription	278

5' End-Labeling	279
3' End-Labeling	279
Probe Purification	279
Probe Storage	280
References	281
13 Practical Nucleic Acid Hybridization	283
Rationale	283
Factors Influencing Hybridization Kinetics and Duplex Stability	284
Temperature	285
Ionic Strength	286
pH	287
Probe Length	287
Probe Concentration	287
G+C Content	288
Mismatching	288
Probe Complexity	289
Viscosity	289
Formamide	290
Urea	290
Hybridization Temperature	291
T _m for Long Probes	291
T _m for Oligonucleotide Probes	292
Hybridization and the Northern Analysis	292
Prehybridization: Filter Preparation	293
Protocol: Prehybridization for Long Probes	293
Probe Denaturation	294
Hybridization	295
Posthybridization Stringency Washes	295
Probe Removal and Rehybridization	296
Protocol: Generic Method for Probe Removal	297
References	298
14 Principles of Detection	301
Rationale	301
Autoradiography	302
Handling of Filter Membranes	304
X-ray Film	305
Safe light	306
Exposure Time	306
Intensifying Screens	307
Fluorography	308
Preflashing Film	309
Type of Cassette	309
Development and Fixation	310

Protocol: Autoradiography	310
Non-isotopic Procedures	312
Biotin	313
Digoxigenin	313
Fluorescence Imaging	314
Direct Enzyme Labeling	314
Detection by Chemiluminescence	314
Substrat for Chemiluminescence	315
Chromogenic Detection Procedures	317
Digital Imaging Systems	318
References	320
15 Quantification of Specific mRNAs by Nuclease Protection	321
Rationale	321
Basic Approach	322
Probe Selection	327
Optimization Suggestions	329
Potential Difficulties	330
Protocol: Transcript Quantification by S1Analysis	332
Protocol: Transcript Quantification by RNase Protection	335
Troubleshooting	337
References	338
16 Analysis of Nuclear RNA	341
Rationale	341
Transcription Rate Assays	343
Relationship to the Study of Steady-State RNA	346
Nuclear Run-off vs. Nuclear Run-on Assay	347
Protocol: Nuclear Run-off Assay	348
Harvesting of Cells and Preparation of Nuclei	349
Alternative Protocol for Preparation of Fragile Nuclei	349
Alternative Protocol for Preparation of Nuclei from Whole Tissue	350
Labeling and Recovery of Transcripts	351
Preparation of Target DNA	353
Preparation of RNA for Hybridization	354
Posthybridization Washes and Detection	355
Protocol: Alternative Procedure for Nuclear Run-off Assay	356
Protocol: Nuclease Protection – Pulse Label Transcription Assay	357
Distinguishing Among the Activities of RNA Polymerases	359
Extraction of Nuclear RNA for Steady-State Analysis	360
Protocol: Direct Isolation of Nuclear RNA	361
Protocol: Preparation of Nuclear RNA from Cells Enriched in Ribonuclease	362
Troubleshoot Nuclear RNA Analysis	363
References	364

17	cDNA Synthesis	367
	Rationale	367
	cDNA Synthesis – An Overview	368
	First-Strand Considerations	369
	Reverse Transcriptase Options	372
	Second-Strand Considerations	375
	Classical Methods	376
	PCR-Based Methods	377
	Protocol: First-Strand cDNA Synthesis	377
	Assessing cDNA Synthesis Efficiency	379
	Cloning cDNA	379
	Ligation Considerations	381
	Enzymes Used for Ligation	383
	Applications	383
	References	384
18	RT-PCR: A Science and an Art Form	385
	Rationale	385
	PCR – An Overview	386
	RT-PCR – General Approach	391
	PCR Carryover Prevention	395
	Laboratory Design	395
	Procedural Methods	396
	Aerosol-Resistant Tips	397
	Uracil-N-Glycosylase	398
	Primer Design	398
	Basic Rules	402
	T _m Considerations	406
	Estimating T _m	406
	Precision T _m Calculations	407
	On-line Resources	407
	Multiplex Primer Design	407
	Optimization Procedures	408
	Thermostable Polymerases	411
	Positive Controls	413
	Negative Controls	413
	Hot Start PCR	413
	Locked Nucleic Acids	415
	Touchdown PCR	416
	Internal Controls	416
	The Word on Transcription Controls	419
	Analysis of PCR Products	420
	RT-PCR Quality Control Points	421
	Non-PCR Methods for Confirming PCR-derived Data	423
	Related Techniques	424

5' RACE PCR	424
5' RLM-RACE	426
3' RACE PCR	426
Nested PCR	428
Long-Range PCR	429
Single Cell PCR	431
Splinkerette PCR	433
The Hunt for Alternative Transcription Start Sites	434
Protocol: First-Strand cDNA Synthesis	435
Protocol: PCR Amplification of cDNA	436
Cloning PCR Products	437
Protocol: A-Tailing of Blunt-End PCR Products	438
Protocol: TA Cloning Ligation Reaction	438
TOPO [®] Cloning	440
Other Amplification Procedures	441
Linear RNA Amplification (Eberwine Process)	441
Strand Displacement Amplification	442
Nucleic Acid Sequence Based Amplification (NASBA)	443
Ligase Chain Reaction	443
References	443
19 Quantitative PCR Techniques	449
Rationale	449
Sensitivity Index	450
Quantitative Approaches	451
Real-Time PCR	453
Real-Time PCR Platforms	456
SYBR Green Assay	457
TaqMan [®] Assay	458
Molecular Beacons	458
LUX [™] Primers	460
Scorpions	461
Melting Curve Analysis	462
Internal Controls	463
Exogenous Controls	463
Control Reaction Formats	467
Negative Control Considerations	471
Competitive PCR: Key Considerations	472
Competitive PCR: Major Steps Involved	477
Alternative Approach: Non-Real-Time Competitive PCR	479
Protocol: Competitive PCR	480
Synthesis of Non-homologous Competitor	480
Synthesis of First-Strand cDNA	481
Competitive PCR (Primary Amplification)	482
Competitive PCR (Secondary Amplification)	484

Image Analysis Considerations	485
Troubleshooting Quantitative PCR Techniques	486
References	488
20 Functional Genomics and Transcript Profiling	491
Rationale	491
Functional Genomics Defined	491
Importance of Functional Genomics Approaches	492
Commonly Used Functional Genomics Approaches	493
Functional Genomics and Classical Molecular Biology	496
21 High-Throughput Analysis of Gene Expression	499
Rationale	499
What is a Microarray?	500
What is a Heat Map?	502
What Microarrays Can Do	504
What Microarrays Cannot Do	504
Major Steps in Microarray Analysis	507
Reference RNA	509
What is a Macroarray?	510
Applications	512
References	512
22 Non-Array Methods for Global Analysis of Gene Expression	515
Rationale	515
Essential Issues	516
Subtractive Method	517
Suppression Subtractive Hybridization (SSH)	517
Troubleshooting	524
Non-Subtractive Methods	526
mRNA Differential Display	526
Protocols	529
References	536
23 RNAi: Take a RISC–Role the Dicer	539
Rationale	539
Essential RNAi Nomenclature	541
RNA Interference – How it Works	542
siRNA Approach	545
shRNA Approach	546
siRNA Delivery Methods into Mammalian Cells	548
miRNA	549
Effective Design of siRNAs	550
RNAi and Alternative Transcript Splicing	552
<i>In Vitro</i> and <i>In Vivo</i> Issues	553

RNAi Validation	555
RT-PCR Approaches	556
Northern Analysis	556
Western Analysis	556
Applications	557
References	557
24 Genomes, Transcriptomes, Proteomes, and Bioinformatics	561
Rationale	561
Essential Nomenclature	563
Genomes and Genomics	564
Transcriptomes and Transcriptomics	565
Proteomes and Proteomics	566
Bioinformatics	571
Search for Genes – Have a BLAST!	571
References	573
25 An RNA Paradigm	575
A Typical Experiment?	577
Sensitivity Issues	580
What to Do Next	580
Where to Turn for Help	582
EPILOGUE	
A Few Pearls of Wisdom	585
Appendix A Maintaining Complete and Accurate Records	591
Appendix B Converting Mass to Moles	593
Appendix C Useful Stock Solutions for the Molecular Biologist	597
Appendix D Phenol Preparation	603
Appendix E Disposal of Ethidium Bromide and SYBR Green Solutions	607
Appendix F DNase I Removal of DNA from an RNA Sample	611
Appendix G RNase Incubation to Remove RNA from a DNA Sample	613
Appendix H Deionization of Formamide, Formaldehyde, and Glyoxal	615
Appendix I Silanizing Centrifuge Tubes and Glassware	617
Appendix J Trypsinization Protocol for Anchorage- Dependent Cells	619
Appendix K Isolation of High-Molecular-Weight DNA by Salting-Out	621

Appendix L	Electrophoresis: Principles, Parameters, and Safety	625
Appendix M	Polyacrylamide Gel Electrophoresis	637
Appendix N	Dot Blot Analysis	641
Appendix O	Centrifugation as a Mainstream Tool for the Molecular Biologist	649
Appendix P	Selected Suppliers of Equipment, Reagents, and Services	655
Appendix Q	Useful SI Units	663
Appendix R	Common Abbreviations	665
Appendix S	Trademark Citations	667
GLOSSARY		671
INDEX		693

This page intentionally left blank

Preface

The Surprising World of RNA

Why study RNA? This question was posed as the first edition of *RNA Methodologies* went to press in 1993. The isolation of RNA continues to be a central though somewhat unappreciated aspect of molecular biology, though much has changed since the third edition of this book was published. The completion of the human genome project has heralded the previously unknown disciplines of functional genomics and proteomics. Researchers now regularly explore the transcriptome, the proteome, the metabolome and, lest we forget, the genome.

To state that RNA is versatile is the understatement of the century. RNA is much more than a DNA-protein go-between. RNA can exhibit catalytic activity, provide binding sites for small molecules for myriad purposes, regulate the expression of other genes, function as the backbone of ribosomes, transport amino acids, and much more. All of these roles support the notion of the widely acclaimed “RNA world hypothesis”, in which RNA, and not DNA, protein, or anything else was the primary primordial information molecule. This is quite plausible because as a single-stranded molecule, RNA has an amazing ability to assume secondary structures and folding properties, imparting functionalities perhaps as diverse as its very nucleotide sequence.

Transcriptional profiling is possible only when high quality RNA is isolated from its biological source, such that it is able to support reverse transcription and any of a variety of downstream applications, including the detection of previously uncharacterized genes, differentially spliced transcripts, or transcripts with multiple start sites. Information of this nature is very important; for example, alternative splicing imparts an added level of vulnerability to mutations and the disease state, meaning that what happens at the level of RNA is often a life or death scenario.

There remains an academic imperative to unify the numerous facets of RNA characterization in a coherent start-to-finish format, though one of the major difficulties toward the realization of that goal is that a rapid succession of new techniques and variants thereof has resulted in confusing technical nomenclature. To make matters worse, not everyone uses the same terminology to describe the same techniques. Quantification these days is of ever increasing importance because of the apparent link between an abnormal abundance of a transcript (too high or too low) and a genetic disease, especially in the context

of the great interest in the use of RNA, and RNA aptamers, for *in vitro* diagnostic applications. While the very sensitive method known as real-time PCR has overshadowed Northern analysis, nuclease protection, and the nuclear runoff assay, these time-honored methods remain important, though comparatively roughhewn, tools for RNA characterization.

Thoughtful and expedient isolation of RNA is equally as important as the assays that follow. The purification of high quality RNA, what this Author affectionately refers to as eRNA (excellent RNA!), from various biological sources is the fundamental starting point for investigations designed to give clearer definition to this aspect of the regulation of gene expression. Whether isolated from cells in culture or directly from whole tissue, only the meticulous handling of RNA will support experiments that will then be used for its study.

This laboratory guide represents a growing collection of tried, tested, and optimized laboratory protocols for the isolation and characterization of eukaryotic RNA, with lesser emphasis on the characterization of prokaryotic transcripts. The more noteworthy additions to the fourth edition include a reorganization of chapters, placing earlier and greater emphasis on protocols. There is also information on a variety of new and improved RT-PCR tricks of the trade, including methods for improving cDNA synthesis, innovative 5' and 3' RACE, subtractive PCR methods, the isolation of RNA from plants, and updated, informative chapters on bioinformatics, high throughput analysis of gene expression, and RNA interference (RNAi), with commentary on the advantages and disadvantages of each. Emphasis is given to how these techniques can be used as stand-alone methods or in unison with other assays for the study of transcriptional and posttranscriptional regulation of eukaryotic genes with the greatest possible sensitivity and resolution.

This text is written for the principal investigator, bench scientist, physician, veterinarian, lab technician, graduate student, undergraduate research assistant, and anyone else capable of performing basic research techniques—there is something in it for everyone. This resource is intended to provide a rationale to assist in the decision-making process for individuals at all levels of sophistication, from the novice to the well-seasoned scientist, and at the same time present realistic alternatives for achieving the same experimental goals. Many of the incorporated notations and hints are based upon personal experience and pave the way for the expedient recovery of RNA and the most judicious use of resources. Day-in and day-out, unsound tactics for RNA characterization result in wasted resources due to an obvious failure to understand the “what” and the “why” from the onset of the study. The best advice that I can offer: always think two steps ahead in an experiment, and reflect upon how the method of RNA isolation will impact the ensuing protocols and interpretation of data.

The recurrent themes herein are the correct way to handle and to assay RNA, and an appropriate level of background information related to the fundamentals of gene expression is likewise provided. These pages demonstrate clearly how a selected technique fits into the grand scheme of nucleic acid

research. While this Author would hope that the text be studied from cover to cover, one may pick and choose salient protocols without loss of continuity. For those readers who are new to the study of the cellular biochemistry from the perspective of RNA analysis, it may be beneficial to first read Chapter 25 (An RNA Paradigm). Collectively, the chapters work together to embellish the RNA story, each presenting clear take-home lessons. The liberal incorporation of flow charts, tables, and representative data likewise facilitate learning and assist in the planning and implementation phases of a project. The investigator is limited only by his own ingenuity.

* * *

The Author acknowledges, with sincere thanks and appreciation, the intellectual encouragement of the many colleagues and friends who, in some way, supported the preparation of this manuscript. The support and PATIENCE of the Author's family are also gratefully acknowledged and are very much appreciated.

Initium sapientiae timor Domini

This page intentionally left blank

1 RNA and the Cellular Biochemistry Revisited

Why study RNA?	1	Coding region	25
What is RNA?	2	Trailer sequence	26
Assembly of polynucleotides	5	Poly(A) tail	26
Types of RNA	8	Organellar mRNAs	28
Transcription and the central dogma	10	mRNA stability, transport, and turnover	29
Important plant and animal models for studying transcription	11	Bicistronic mRNAs	30
Promoters and regulatory elements	12	Prokaryotic mRNAs	30
Gene and genome organization affect transcription	14	mRNA sequence and structure affect translation	31
RNA polymerases and the products of transcription	18	Levels of gene regulation	32
Messenger RNA	22	Alternative splicing of mRNA from a single genetic locus	34
Topology of a typical mRNA molecule	23	<i>Trans</i> -splicing: mRNA repair	35
5' Cap	23	Overview of small RNAs	36
Leader sequence	25	miRNA regulation of gene expression	37
		References	38

Why study RNA?

All cell and tissue functions are ultimately governed by gene expression. Consequently, the reasons for electing to study the modulation of RNA as at least one parameter of the cellular biochemistry may be as diverse as the intracellular RNA population itself. Generally speaking, the characterization of RNA is almost always related to transcription (i.e., gene expression) questions being asked in the context of a particular scientific inquiry, and most often revolves around measuring the dynamic abundance level of one or several transcripts.

The goals in any experimental design involving RNA generally revolve around one or more fundamental themes, including but not limited to the following:

1. Measurement of the steady-state¹ abundance of cellular transcripts. This is the most commonly studied parameter of gene expression primarily because of the ease with

¹The final accumulation of RNA in the cell, or in a subcellular compartment such as the nucleus or the cytoplasm, is referred to as its steady-state level and is the net result of the synthesis, stability, and degradation of the highly variegated products of transcription.

which RNA can be isolated. Quantitative and qualitative profiles of a population of RNAs can then be generated by contemporary methods, such as PCR-based approaches, or through the use of time-honored techniques such as Northern analysis or nuclease protection.

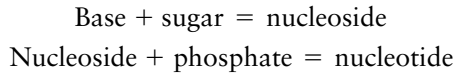
2. Measurement of the *rate* of transcription of gene sequences or the pathways of RNA processing. This may be deduced, at least in part, by the nuclear run-on assay in which radiolabeled ribonucleotide precursors are incorporated into nascent transcripts, which occurs in direct proportion to the abundance of each RNA being transcribed. When used in conjunction with Northern analysis or other methods that examine steady-state RNA levels, the regulation of genes can often be assigned as transcriptional or due to posttranscriptional events.
3. Identification of the transcription start site (TSS). Historically, mapping of RNA molecules, including the 5' end, the 3' end, and the size and location of introns, was accomplished via the nuclease protection assay, as described in Chapter 15. Now, however, transcript mapping is now almost always performed by rapid amplification of cDNA ends (5'- and 3' RACE; see Chapter 18). Because a single genetic locus can produce multiple RNAs, each with a different TSS, and often in a tissue-specific manner, this type of analysis is extremely useful and in widespread use.
4. Synthesis of complementary DNA (cDNA). Unstable, single-stranded mRNA can serve as the template for the *in vitro* synthesis of very stable single- or double-stranded cDNA molecules. Among the more common reasons for performing cDNA synthesis is the amplification of the sequence by PCR, often for some "quantitative purpose", for transcript mapping purposes, for direct ligation into a vector for sequencing or for expression of the encoded protein, for the physical separation of two or more cDNA species, or for the synthesis of an entire cDNA library² which can be propagated for long-term storage and analysis. In so doing, one constructs a permanent record of the cellular biochemistry at the moment of cell lysis or tissue disruption. Historically, the synthesis of highly representative cDNA is one of the most important, yet technically challenging, methodologies in the molecular biology laboratory.
5. *In vitro* translation of purified messenger RNA (mRNA). The resulting polypeptide may be further characterized by immunoprecipitation or Western analysis. Cell-free translation represents at least one method for the identification of specific transcripts: by providing the raw materials needed to support translation, one is able to demonstrate that a transcript of putative identity is able to support the synthesis of the cognate peptide. For example, this approach could be used to demonstrate that two transcripts from the same genetic locus with alternative transcription start sites are, in fact, able to direct the synthesis of identical or closely related proteins. In applications such as rational drug design, *in vitro* translation is also helpful because understanding the three dimensional architecture of a protein, and its wild type or mutated function(s), may suggest novel applications in the area of functional proteomics.

What is RNA?

RNA is a long, unbranched polymer of ribonucleoside monophosphate moieties joined together by phosphodiester linkages, and both eukaryotic and prokaryotic RNAs are essentially single-stranded molecules. The unassembled

²Older literature often refers to a cDNA library as a "clone bank".

monomers of both RNA and DNA are called nucleotides. These building blocks consist of three key components: a pentose (five-carbon sugar), at least one phosphate group (nucleotides may contain as many as three phosphate groups), and a nitrogenous base (Fig. 1.1). A nitrogenous base joined to a pentose sugar is known as a nucleoside. When a phosphate group is added, the composite, a phosphate ester of the nucleoside, is referred to as a nucleotide.



The key chemical difference between RNA and DNA is the presence the five-carbon sugar ribose, in which a hydroxyl group ($-\text{OH}$) is joined to the 2' carbon of the ribose sugar whereas the absence of this $-\text{OH}$ group in DNA is the underlying basis of the name of the sugar deoxyribose. In addition, one finds the base uracil in RNA, substituted in DNA by the closely pyrimidine thymine³, though it is possible to find nucleotides containing uracil in DNA in certain situations. More precisely, RNA is assembled from ribonucleotide precursors and DNA is assembled from deoxyribonucleotide precursors. Hence, RNA is so-named because of the ribose sugar it contains, just as DNA is named from its constituent 2'-deoxyribose sugar. The principal features and essential nomenclature of these building blocks are summarized in Tables 1.1 and 1.2.

Nitrogenous bases and the pentose sugar components of nucleosides are both cyclic. To avoid confusion when referring to the constituent atoms of the sugar or the base that constitute a particular nucleotide, a special numbering

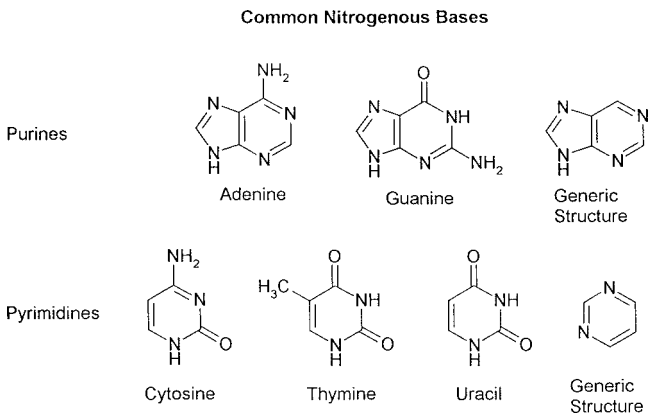


Figure 1.1 The identity of a nucleotide is defined by the base that is attached to the 1' carbon. In practice, the nucleotides that make up an RNA or DNA molecule are represented by the standard one-letter abbreviation for the base each contains: adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U).

³Chemically, thymine is methyluracil.

Table 1.1 Comparative Nucleotide Structure

	RNA nucleotides	DNA nucleotides
Key nucleotide components	Five-carbon ribose Phosphate group(s) Nitrogenous base	Five-carbon deoxyribose Phosphate group(s) Nitrogenous base
Common nitrogenous bases		
Purines	Adenine Guanine	Adenine Guanine
Pyrimidines	Cytosine Uracil	Cytosine Thymine

Table 1.2 Essential Base, Nucleoside, and Nucleotide Nomenclature

	Base	Nucleoside	Nucleotide triphosphate precursors
RNA	Adenine (A)	Adenosine	Adenosine-5'-triphosphate (ATP)
	Cytosine (C)	Cytidine	Cytidine-5'-triphosphate (CTP)
	Guanine (G)	Guanosine	Guanosine-5'-triphosphate (GTP)
	Uracil (U)	Uridine	Uridine-5'-triphosphate (UTP)
DNA	Adenine (A)	2'-deoxyadenosine	2'-deoxyadenosine-5'-triphosphate (dATP)
	Cytosine (C)	2'-deoxycytidine	2'-deoxycytidine-5'-triphosphate (dCTP)
	Guanine (G)	2'-deoxyguanosine	2'-deoxyguanosine-5'-triphosphate (dGTP)
	Thymine (T)	2'-deoxythymidine	2'-deoxythymidine-5'-triphosphate (dTTP)

Nucleosides consist of a base and a sugar only. A nucleoside is promoted to a nucleotide upon addition of at least one phosphate group.

system has been devised. By convention, the numbering system for the carbon and nitrogen atoms that make up the bases is 1, 2, 3, and so forth, while the numbering system for the constituent carbon atoms of the sugar (ribose or deoxyribose) is 1', 2', 3', 4', and 5'.

The ribonucleoside triphosphates may be collectively (and sometimes singularly) referred to as NTP; in various molecular biology protocols the symbol NTP refers to an equimolar cocktail of ATP, CTP, GTP, and UTP. Similarly, the deoxy- form of a nucleotide is denoted by the placement of a lower case "d" preceding the nucleotide triphosphate, as in dATP, dCTP, dGTP, and dTTP, and the symbol dNTP refers to an equimolar cocktail of the four deoxynucleoside triphosphates in protocols capable of supporting the *de novo* synthesis of cDNA or PCR products. It is the triphosphate form of each of these nucleotides that is recognized by RNA polymerase and DNA polymerase, respectively, thereby serving as precursors in nucleic acid synthesis. The phosphate nearest the nucleoside moiety, that is, closest to the sugar, is known as the α phosphate,

Essential Nucleotide Structure

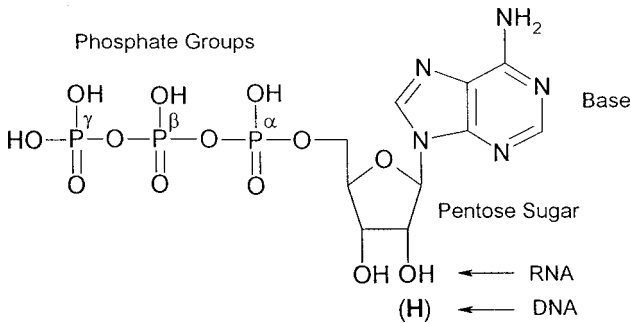


Figure 1.2 Adenosine-5'-triphosphate (ATP). The three constituent phosphate groups are designated α , β , and γ based on the proximity of each group to the nucleoside (base + sugar) component of the molecule. The replacement of the 2'-OH with H would convert this molecule to a deoxynucleotide.

followed by the β phosphate, followed by the γ phosphate, which is furthest from the nucleoside (Fig. 1.2). During nucleic acid polymerization, the β and γ phosphates are cleaved from the nucleotide, and the resulting single-phosphate nucleotide is then incorporated into the nascent polynucleotide chain.

Assembly of polynucleotides

The synthesis of RNA is mediated by the activity of enzymes known as RNA polymerases while DNA is synthesized via the enzymatic activity of DNA polymerases. Thus, any enzyme with an associated polymerase activity is capable of synthesizing nucleic acid molecules from nucleotide precursors. A nucleic acid molecule is the result of linking nucleotides together by phosphodiester bonds. The formation of these bonds involves the hydrophilic attack by the 3' hydroxyl group of the last nucleotide added to the nascent polynucleotide on the 5' phosphate group of the incoming nucleotide (Fig. 1.3).

Assuming that the intracellular biochemistry can support initiation and elongation of RNA molecules, there are two fundamental requirements that must be fulfilled and maintained *in vivo* (and *in vitro*) to support continued nucleic acid polymerization:

1. There must be a template strand to direct the polymerase-mediated insertion of the correct (complementary) nucleotide into the nascent chain. This occurs predictably, according to the conventions set down in Chargaff's Rule (Zamenhof *et al.*, 1952), which succinctly states that adenine ordinarily base-pairs with thymine or uracil through the formation of two hydrogen bonds (A::T, A::U) and that guanine ordinarily base-pairs with cytosine through three hydrogen bonds (G:::C).
2. For initiation and elongation, there must be a free 3'-OH to which the next nucleotide in the chain can be joined via a phosphodiester linkage. Thus, the entire process of transcription *in vivo* and *in vitro* requires some type of primer manifesting the

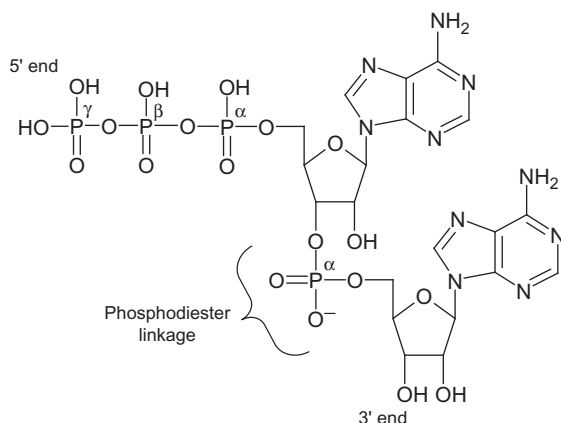


Figure 1.3 The dinucleotide that results from the formation of the first phosphodiester linkage has structurally different ends, namely an unpaired phosphate (the 5' end) and a hydroxyl group (the 3' end). The structural differences at the 5'- and 3'-ends are maintained regardless of the number of nucleotides that are joined together.

requisite 3'-OH. The same is true for DNA synthesis⁴. Most of the enzymes used in molecular cloning that exhibit polymerase activity have nearly identical template and 3'-OH primer requirements.

Polynucleotide elongation involves the incorporation of a nucleoside monophosphate, and the formation of this linkage is accompanied by the release of pyrophosphate. The synthesis of nucleic acids, both RNA and DNA, always proceeds from the 5' direction toward the 3' direction, resulting in a polynucleotide that manifests a consistent pattern of 5'→3' linkages between adjacent nucleotides. Elongation is frequently referred to as the 5'→3' polymerase activity associated with the enzyme.

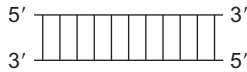
Upon completion, nucleic acid molecules are assembled in such a way that:

1. The ends of the molecule are structurally different from one another. The first nucleotide of the molecule has an uninvolved 5' (tri)phosphate, constituting the so-called 5' end of the molecule. The last nucleotide that was added manifests a free 3' hydroxyl group, and this is known as the 3' end of the molecule.
2. The backbone of the molecule consists of an alternating series of sugar and phosphate groups. Known as the phosphodiester backbone, or simply the backbone, of the molecule, it imparts a net negative charge to the molecule by virtue of its constituent phosphate groups.
3. The base associated with each nucleotide protrudes away from the backbone of the molecule. This stereochemistry makes the bases very accessible for hydrogen bonding (base

⁴The only DNA polymerase capable of adding nucleotides *without* template information is the unusual enzyme terminal deoxynucleotidyl transferase, commonly known as terminal transferase. This enzyme has broad applications in the area of cDNA synthesis as well as certain forms of 5'-RACE. These special cases are discussed in detail in Chapters 17 and 18. A few other enzymes show *limited* terminal transferase activity, and these are likewise discussed in subsequent chapters.

pairing) to a complementary polynucleotide sequence which, *in vivo*, is how double-stranded molecules form and *in vitro*, is the very heart of molecular hybridization.

The nitrogenous bases found in nucleotides are categorized as either purines (adenine and guanine) or pyrimidines (cytosine, thymine, and uracil), both of which are flat aromatic molecules. The specificity of base pairing (purine with pyrimidine) is maintained by the preference of adenosine and cytosine for the amino over the imino form and the preference of guanine, thymine, and uracil for the keto rather than the enol form. In other words, what is commonly known as “Watson–Crick” base-pairing is predicated on the bases involved being in their preferred tautomeric forms. Hydrogen bonds, which are highly directional, form between complementary bases occurs when an electropositive hydrogen atom is attracted to an electronegative atom such as oxygen or nitrogen. Because of the manner in which bases protrude from their respective phosphodiester backbones, base pairing or hybridization of complementary strands in an *antiparallel* (opposite) configuration is strongly favored. This is true for all double-stranded molecules: dsDNA, dsRNA, and DNA:RNA hybrids. Thus, the 5′ end of one strand is opposite the 3′ end of the complementary strand to which it is base-paired and often represented as shown here.



The obvious structural differences at the 5′ and 3′ ends of a molecule support a convention by which one may unambiguously refer to the position of any topological feature of a nucleic acid molecule in relation to any other topological feature:

Upstream means that the salient structure or feature is closer to or in the direction of the 5′ end of the molecule, relative to some other point of reference.

Downstream means that the salient structure or feature is closer to or in the direction of the 3′ end of the molecule, relative to some other point of reference.

For the sake of simplicity, upstream and downstream may be used to mean “in the opposite direction of expression” and “in the direction of expression”, respectively. This nomenclature may be especially useful when describing the salient features of a double-stranded nucleic acid molecule, in discussions pertaining to either the structure or the expression of a gene and, in particular, for the purpose of primer design to support PCR.

The actual base sequence of a polyribonucleotide is known as the primary (1°) structure of the RNA. There is a tremendous proclivity for sequences within a single RNA molecule to exhibit intramolecular base pairing, perhaps better known as secondary (2°) structure. The variety of possible interactions within the phosphodiester backbone are often described using such colorful nomenclature as RNA hairpin, stem, interior loop, bulge loop, multibranching loop, and pseudoknot. Higher order three-dimensional folding, the so-called tertiary (3°) structure which RNA molecules exhibit, is best described as the collection of 2° structural elements required by an RNA molecule to perform its biological function. Much has been suggested about the role of folding by

careful study of transfer RNAs, the classical example of intramolecular base-pairing *par excellence*. Often, much of the 2° and 3° structure of tRNA may be attributed to the formation of non-canonical⁵ base-pairs. More contemporary studies have demonstrated that mRNA also assumes varying degrees of transient 2° and 3° structure which, in no small measure, influence its function in the cytoplasm. For most laboratory applications, higher-order folding must often be disrupted, as described below, before an assay with a quantitative component can be performed using a purified RNA. Failure to do so generally has a severe negative impact on accurate quantitative profiling of the sample.

Types of RNA

Transcription results in the production of RNA molecules, often generically referred to as *transcripts*. Traditionally, the transcripts observed within a cell were broadly classified as ribosomal RNA (rRNA), transfer RNA (tRNA), heterogeneous nuclear RNA (hnRNA), or messenger RNA (mRNA), as well as a collection of small RNAs of previously unknown function. Now, however, one must include the extremely diverse population of miRNAs and other RNAs which are of immense interest in the study of the regulation of gene expression (Table 1.3). Each category of RNA, which is synthesized by a different type of RNA polymerase, performs a different function in the cell. These highly diverse populations of RNA are not represented in equal amounts in the cell and the relative amount of each is directly related to the physiology of the cell.

rRNA is the most abundant RNA component in the cell. In prokaryotic cells the major rRNA species are the 23S rRNA, 16S rRNA, and 5S rRNA. The eukaryotic counterparts are identified as the 28S rRNA, 18S rRNA, and 5S rRNA, as well as a fourth ribosomal transcript, the 5.8S rRNA. These molecules form the scaffolding of ribosomes, which become translationally competent when decorated with myriad ribosomal proteins. At present there are 55 known prokaryotic ribosomal proteins and 82 known eukaryotic ribosomal proteins (Watson *et al.*, 2008). Not all ribosomes are functional at any given time, and the existence of a pool of transiently inactive ribosomes is itself a regulator of gene expression. The super abundance of rRNA is often exploited as both an RNA mass loading control (Chapter 19) as well as internal molecular weight markers for electrophoresis (Chapter 9).

tRNA is responsible for the transportation of amino acids to the ribosome to support *de novo* protein synthesis. These amino acid shuttle molecules are small, ordinarily ranging from 74 to 95 bases. When carrying an amino acid covalently linked to its 3' end, a tRNA is said to be “charged”. Within the coding region of mRNA, each group of three nucleotides constitutes a codon; placement of the correct amino acid into the nascent polypeptide depends on

⁵Canonical base pairs: G-C and A-U; non-canonical base pairs: A-C, G-U, C-U, U-U, G-G, A-Ψ, G-Ψ, A·A·U trimers, and others (Ψ = pseudouridine). An excellent database containing known non-canonical base pairs involving RNA is maintained by Dr. George Fox at http://prion.bchs.uh.edu/bp_type/, the publication reference for which is Nagaswamy *et al.*, 2002.

Table 1.3 RNA Types and Functions

RNA type	Symbol	Basic function	Prokaryotic	Eukaryotic
Ribosomal RNA	rRNA	Forms back bone of the ribosomal subunits	Yes	Yes
Transfer RNA	tRNA	Transports amino acids to the ribosome to support translation	Yes	Yes
Messenger RNA	mRNA	Template for the synthesis of proteins	Yes	Yes
Heterogeneous nuclear RNA	hnRNA	Large unspliced precursor of mRNA (pre-mRNA)	No	Yes
Small nuclear RNA	snRNA	Facilitates splicing of hnRNA into mature, functional mRNA	No	Yes
Small nucleolar RNA	snoRNA	Processing of immature rRNA transcripts in the nucleolus	No	Yes
Small cytoplasmic RNA	scRNA	Facilitates protein trafficking and secretion	Yes	Yes
Micro RNA	miRNA	Short antisense RNAs that participate in the regulation of gene expression	No	Yes
RNase P RNA	–	Catalytic RNA component of the enzyme/RNA complex that processes tRNA molecules	Yes	No
Telomerase RNA	–	RNA component of the enzyme/RNA complex that repairs chromosome telomeres	No	Yes

recognition of the mRNA codon by a complementary trinucleotide motif carried on one arm of the tRNA and known as an anticodon. The tRNA anticodon base pairs to the codon within the ribosome, thereby supporting protein elongation (for review, see Lewin, 2008). While not as abundant as rRNA, the smaller tRNA species play a central role in translation.

mRNA is the most diverse of all the transcripts. Ironically, even though mRNA is by far the least abundant of all transcript types, it is the mRNA that drives the phenotype of the cell. mRNA alone directs the synthesis of proteins through the use of the cellular translation apparatus consisting in large measure of rRNA and tRNA. Specific mRNA species vary widely in number and abundance in the cell. Some mRNAs are present in hundreds of copies per cell while others still are present only a few copies per cell; this aspect of the RNA profile of the cell can be problematic because very low abundance transcripts are often difficult to detect even with sensitive contemporary techniques such as real-time PCR. Functional mRNA in eukaryotic cells is derived from spliced and modified nuclear transcripts, known collectively as hnRNA, and it is the mRNA which is most often of direct experimental interest to the molecular biologist.

Transcription and the central dogma

According to the central dogma⁶ of molecular biology, the expression of hereditary information flows from genomic sequences (DNA), through a messenger RNA (mRNA) intermediate, to ultimate phenotypic manifestation in the form of a functional polypeptide (Fig. 1.4). Whereas this design mirrors what occurs in both prokaryotic and eukaryotic cells, certain “violations” have been observed in nature: (1) accompanying the discovery of the retroviral enzyme reverse transcriptase⁷ (Baltimore, 1970; Temin and Mizutani, 1970), by which RNA may serve as the template for the synthesis of DNA; and (2) the discovery of RNA editing (Benne, *et al.*, 1986), also known as RNA sequence alteration, in which a transcribed sequence is subject to change.

Nearly a century and a half have past since the first notion of what we now know as a gene was proposed by Father Gregor Mendel, a Roman Catholic priest living in Brno, then part of Austria. The results of the numerous genetic crosses that Mendel set up using garden peas caused him to recognize that what he called “factors” were being passed from one generation to the next. The nature of these factors, however, was completely unknown, and perhaps even more remote were the mechanics of the expression of these factors. The

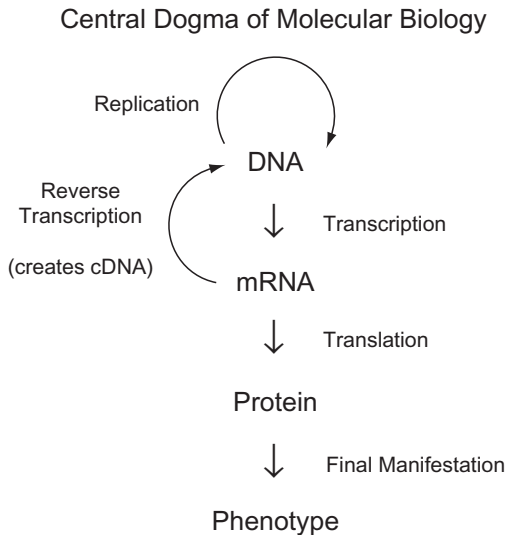


Figure 1.4 The central dogma of molecular biology. The process of transcription produces mRNA while the process of translation produces protein. Replication, the process by which DNA is duplicated, occurs during S phase in the eukaryotic cell cycle. Complementary DNA (cDNA) is an *in vitro* phenomenon used to assay transcriptional activity.

⁶Coined by Francis Crick in 1957 in his paper entitled “On Protein Synthesis”, presented to a symposium of the Society for Experimental Biology.

⁷RNA-dependent DNA polymerase.

last 60 years in particular have seen unparalleled scientific inquiry and discovery in the areas of the nature of the genetic material, the organization of the genetic material in the cell, the wonders of gene expression, and the regulation thereof.

Transcription is that process by which a single-stranded RNA molecule is synthesized from a specific chromosome locus; this is the first of several steps in what is commonly referred to as RNA biogenesis. Transcription occurs in the nucleus (and mitochondria and chloroplasts) of eukaryotic cells, and in the common cellular compartment in prokaryotic cells. All phases of transcription are subject to variation and are potential control points in the regulation of gene expression. A transcriptional unit, therefore, is a DNA sequence that manifests appropriate signals for the initiation and termination of transcription and is capable of supporting the synthesis of a primary RNA transcript. The process of transcription is so named because the transfer of information from DNA to RNA is in the same language, namely the language of nucleic acids. In contrast, the process known as *translation* is so named because nucleic acid instructions in the form of mRNA are used to direct the assembly of a primary polypeptide from amino acid precursors: the nucleic acid instructions are executed in the language of proteins. The ribosome is the organelle of polypeptide synthesis in all cells, and each ribosome independently directs the sequential linkage of amino acids as the associated mRNA is interpreted. Upon completion of translation eukaryotic proteins are often modified, sorted, packaged, and directed to their proper subcellular location by the Golgi apparatus, while prokaryotic and other eukaryotic proteins are often under the influence of various scRNA species.

Important plant and animal models for studying transcription

Current understanding of transcriptional control of gene expression in animals is the result of classical research involving the use of unassuming animals such as *Drosophila*, sea urchins, rats, mice, and the nematode *Caenorhabditis elegans*. Animal cell culture models, including the Chinese hamster ovary cell line CHO-K1 and the Swiss mouse fibroblast cell line NIH 3T3, continue their long-standing role as standard cell lines for the discernment of the esoteric details of transcription and transcript processing. Likewise, in plants, the genomes of *Arabidopsis* (*Arabidopsis thaliana*), rice (*Oryza sativa*), and poplar (*Populus tricarpa*) have been completely sequenced and many other sequencing projects, including apple and peach, are nearing completion. Because *Arabidopsis* was the first plant whose genome was completely sequenced, it has served as a model plant for gene organization, structure, and expression. The ready availability of genome information for many plants, animals, and microorganisms has greatly facilitated the study of functional genomics, that is, the expression of the gene sequences encoded therein.

Promoters and regulatory elements

Transcription is mediated by enzymes known as RNA polymerases. These enzymes, in conjunction with myriad proteins known as transcription factors, recognize very specific and highly conserved *promoter*, or initiation, sequences within the enormous complexity of genomic DNA. Promoters are spatially associated with the structural portion (body) of a gene (Fig. 1.5) and consist of several recognizable nucleotide motifs. These sequences are also commonly known as consensus sequences, a term used to describe the most commonly observed pattern of nucleotides at a particular location. For example, the symbol $T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$ indicates that thymine is the first base associated with this consensus motif 80% of the time, and so forth. The exact sequence and precise geometry of these regulatory elements can either promote or prevent the onset of transcription, and do so with varying degrees of efficiency.

A promoter element that is located 5', or upstream, from the transcription start site (TSS) is indicated with a "minus" sign in front of the actual nucleotide distance from the TSS. By convention, the first transcribed nucleotide is designated as +1, and any other nucleotides or features located 3', or downstream, from the TSS are likewise designated with a "plus" sign placed in front of the actual nucleotide distance. Knowledge of promoter consensus sequence function is due largely to experiments involving standard DNA cloning techniques, site-directed mutagenesis, DNA sequencing, and *in silico* analysis.

In prokaryotic systems, the elements of the promoter region include the so-called -10 sequence, formerly known as the Pribnow box, consisting of the consensus sequence TATAAT, and another conserved region located upstream is known as the -35 sequence (TTGACA). In some organisms, an AT-rich domain (the UP element) is also observed further upstream. The spacing between the -10 sequence and the -35 sequence is tightly regulated, with 17 bp

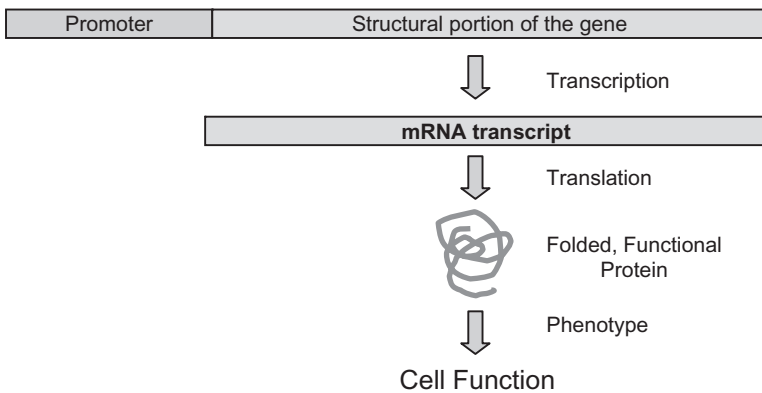


Figure 1.5 Genes, encoding mRNA which in turn encode proteins, are under the direct influence of a regulatory element known as a promoter.

being optimal, and variations in the length of the region between these two elements can destabilize RNA polymerase interaction with the promoter.

In eukaryotic cells, promoters associated with nuclear genes are variable in structure; these variations are due unquestionably to the presence of multiple nuclear RNA polymerases. In particular, promoters recognized by RNA polymerase II (discussed below) display similar sequence homology with and, arguably, resemble prokaryotic gene promoters in many instances. The eukaryotic promoter counterpart is known as the “TATA box”, formerly known as the Hogness box, and so named because of the prevalence of the highly conserved TATAA motif. Point mutations involving any of these five bases strongly downregulate the function of the associated promoter. While at one time it was thought that all eukaryotic promoters manifest a TATA box, this is now known to be untrue. Instead, these rather prevalent TATA-less promoters are characterized by what has been termed a downstream promoter element, or DPE, which is observed approximately 30 bases downstream (+30) from the transcription start site. In addition to the TATA- or DPE structures, another promoter motif is the “CAAT box”, found in several but not all promoters, and so named because of the conservation of its sequence. When present in eukaryotic promoters, the TATA box is usually centered at -30 and the CAAT box appears around -75 , though the CAAT box has been shown to function quite effectively much further upstream. These elements appear to control initial binding of the RNA polymerase and promoter efficiency, respectively. A third frequently observed promoter element is the sequence $(GGGCGG)_n$, known as the G-box element or simply as the GC box. Present in one or more copies, this GC-rich region is generally observed between -90 and -120 within the promoter region. Another short element, the octamer ATGCAAAT and its inverted complement (ATTTGCAT) have also been identified as promoter-associated sequences that influence the efficiency of transcription by RNA polymerase II; these sequences play a role in transcription factor-binding and initiation of transcription. Interestingly, it appears that there is no one component or organization that is shared by all promoters, though the particular permutation of promoter elements and distances between them is recognizable as a transcription initiation regulator. Succinctly, by comparison with transcription in prokaryotic cells, the initiation of eukaryotic transcription requires the presence of numerous transcription factors, coactivators, and transcription activator proteins that bind to these *cis*-acting components which, collectively, make up a promoter. It has been argued that the role of transcription factors binding to gene promoters is to recruit RNA polymerase to that site, though a better description may be the capturing of RNA polymerase so as to initiate transcription. Rather than being thought of as merely an on-off switch associated with a particular gene, a promoter functions more like a thermostat that increases (upregulates) and decreases (downregulates) the expression of a gene in response to the prevailing local conditions acting upon a cell.

Eukaryotic promoters do not necessarily function alone. In many cases, transcription in eukaryotic cells can be influenced profoundly by the presence of an

enhancer sequence, the function of which appears to be the stimulation of transcription initiation. The precise location and orientation of enhancer sequences, with respect to the gene promoter, varies from one gene to the next. Some genomic sequences, such as the immunoglobulin genes, even carry enhancers within the structural portion of the gene itself. Removal of enhancer sequences can reduce the transcriptional efficiency of a locus normally under the influence of an enhancer sequence, as can the binding of repressor proteins to protagonist DNA sequences known as silencers. *In vitro* transcription of genes that are not naturally associated with an enhancer element can be increased tremendously if an enhancer is ligated to the DNA, usually in no particular orientation, and often hundreds, if not thousands of base pairs away from the transcription start site. *In vivo*, a translocation event that brings a promoter and a gene into proximity can result in inappropriate expression of the gene, often with potentially catastrophic consequences, as in the case of Burkitt's lymphoma (Taub *et al.*, 1982). The transcriptional influence of upstream and downstream enhancer and silencer sequences on gene promoters is well documented. It appears that enhancers perform their function(s) by increasing the concentration of transcription activator proteins in the vicinity of the promoter.

During transcription, both strands of the gene being transcribed have different names and different roles. The strand that actually serves as the template upon which RNA is polymerized is properly referred to as the *template strand*. The other strand, which does not act in a template capacity, is called the *coding strand*⁸. When publishing a gene sequence, the convention is to show the sequence of the coding strand, written 5' to 3', from left to right. The implication is that the other strand, the template strand, is base-paired to and lying 3' to 5', or antiparallel, with respect to the coding strand. The template strand is so named because the precise sequence of nucleotides inserted into the nascent RNA transcript is determined by and complementary to the template strand nucleotide sequence. It is important to realize that the coding strand and template strand may switch roles depending upon the placement of transcriptional promoter sequences. The quintessential example of this phenomenon *in vitro* is the cloning of a double-stranded DNA between two different RNA promoters, often the bacteriophage polymerase promoters SP6, T3, or T7. Constructions such as these are frequently employed to accommodate *in vitro* transcription of large amounts of sense and/or antisense RNA for use as nucleic acid probes (Chapter 12) or for RNAi applications (Chapter 23).

Gene and genome organization affect transcription

In order to understand the significance of the products of transcription, it is first essential to understand the organization of the genes themselves. The typical

⁸The coding strand is also known in some circles as the sense strand, while the template strand may be referred to as the antisense strand.

prokaryotic genome exhibits little extraneous baggage. Frequently, genes that encode proteins associated with a common metabolic pathway are clustered together, as suggested by the operon model (Jacob and Monod, 1961). The *lac* operon, the gene products of which facilitate the metabolism of lactose as a carbon source in bacteria, is but one extremely well-characterized example. The RNA molecule that results from the transcription of an operon is usually polycistronic, meaning that more than one polypeptide is encoded in a single RNA transcript. The coding information within a polycistronic mRNA for *each* polypeptide is contiguous: there are no interruptions in the coding sequences by extraneous or non-coding information. This design favors maximum efficiency of energy resources in unicellular organisms.



In fact, the kinetics of prokaryotic gene expression are so rapid that bacterial mRNA is usually being transcribed, undergoing translation, and being degraded simultaneously. These intrinsic factors have in the past frustrated valiant attempts to clone or otherwise characterize prokaryotic mRNA. While a great many improvements that favor the isolation of useful RNA from both Gram-negative and Gram-positive bacteria have been made, and some of these innovations are available in kit form, the isolation of representative prokaryotic RNA remains something of a challenge.

In contrast to prokaryotic systems, nearly all eukaryotic mRNAs are monocistronic. Although a single polypeptide species results from the translation of a particular monocistronic eukaryotic mRNA molecule, that same mRNA is subject to repeated translation as long as the transcript remains biologically and chemically competent. To maximize the translation potential, an mRNA transcript is often engaged by several ribosomes that are all involved in simultaneous, orderly translation of a single transcript. Such a cluster of ribosomes attached to a single mRNA molecule is known as a polysome. Polysomes are observed both in prokaryotic and eukaryotic cells, though eukaryotic polysomes, with 7–8 ribosomes per polysome, tend to be smaller than their prokaryotic counterparts. The entirety of mRNAs so engaged in a cell at any given moment is known as the polysome fraction and which can be used to assess the translational competence of a cell under a defined set of experimental conditions. Succinctly, a large number of polypeptide molecules can be manufactured from a single RNA molecule. In the cytoplasm both polysomal and non-polysomal mRNAs exist as mRNP complexes. Although the majority of polysomes have been shown to be attached to the cytoskeleton (Lenk *et al.*, 1977; Davies *et al.*, 1991), no absolute requirement for this association and concomitant translation has been shown.

Close examination of eukaryotic genes reveals that for a vast majority of genes there are considerably more nucleotides within a particular locus than are necessary to direct the synthesis of the corresponding polypeptide, that is, the DNA sequence and the amino acid sequence are not colinear over the span

of the locus. This size differential can also be observed at the level of the mature mRNA in the cytoplasm, which is usually quite a bit shorter than the DNA sequence from whence it was derived. Upon further scrutiny, this discrepancy can be resolved at the level of the organization of the gene itself, the sequences within which fall into one of two categories:

1. *Exons* are regions of DNA that are represented in the corresponding mature cytoplasmic RNA. Exons may or may not have a peptide coding function.
2. *Introns* are regions of DNA that are transcribed but generally are not represented in the corresponding mRNA. Introns are systematically spliced out of the primary RNA transcript. This results in the joining of exon sequences that lie adjacent to introns. While introns usually do not direct polypeptide synthesis, there are several noteworthy exceptions (for review, see Farrell and Bassett, 2007).



The number and length of exons and introns associated with a gene are highly variable depending upon locus and this variability even pertains to loci that are highly conserved across evolutionary time. By comparison with introns which can be several thousand base pairs in length, exons tend to be rather short, each encoding fewer than 100 amino acids in most organisms. In many cases the high conservation of sequence in one or more exons of a gene has been directly responsible for the isolation of a related gene from a different organism. There is no significant homology to be found among intron sequences, and introns often exhibit multiple termination codons in all reading frames. This is not entirely unexpected because the non-coding function of introns favors the accumulation of mutations that might otherwise be lethal were they to occur within an exon sequence. In some unusual cases, genes such as human β -interferon lack intron sequences completely (Tavernier *et al.*, 1981). Examination of the splicing junctions of introns, however, reveals a strict conservation of two dinucleotide consensus sequences (Breathnach and Chambon, 1981; Mount, 1982) contained entirely within the intron. Proceeding from the 5' end of the RNA, introns are found to begin with a GU dinucleotide and end with an AG dinucleotide⁹; this phenomenon is believed to occur 100% of the time in higher eukaryotes.



The nucleotides immediately adjacent to both sides of the dinucleotide intron boundaries are also conserved to an extent (65–75%) and a point mutation at a splice site generally results in the inactivation of that site. This phenomenon is the basis of a method for discerning exon sequences known as exon trapping (Duyk *et al.*, 1990) in which a putative exon-containing sequence

⁹The so-called GT-AG rule, describing exon-intron splice sites, refers to the DNA sequence. The primary transcript itself manifests GU at the 5' end of an intron.

is cloned into a specialized vector that consists of an intron flanked by two known exons; an exon in the cloned DNA, when ligated into the vector intron will result in a longer transcript that can be assayed by Northern analysis. In some cases, splice-site mutations can result in the production of an aberrant mRNA through the use of an alternative splice site, often located within the intron (Triesman *et al.*, 1982), as in certain β -thalassemic individuals. Whereas this interrupted gene organization is probably common to all higher eukaryotes, the consensus phenomenon does not apply to mitochondrial or chloroplast loci, nor to yeast tRNA genes (Lewin, 2008).

The mechanics of intron removal and exon ligation are mediated in part by a highly conserved family of small nuclear RNAs (snRNA). These molecules exist as the RNA-protein complexes, known as U1, U2, U4, U5, and U6, are confined to the nucleus where they are generically referred to as small nuclear ribonucleoproteins (snRNPs, or snurps). Methodologies for efficiently characterizing these essential splicing cofactors were first described by Hamm and Mattaj (1990), and snRNPs are now known to form enormous complexes known as spliceosomes, which are known to mediate pre-mRNA splicing. A similar, abundant group of small cytoplasmic RNAs (scRNA) are found in the eukaryotic cytoplasm. As with snRNAs, scRNA molecules are known to exist as RNA-protein complexes (scRNP, or scyrps). The role of the scRNA molecules in regulating the synthesis, sorting, and secretion of proteins as well as possible mRNA degradation, remains an area of great interest. Yet another class of small nucleolar RNAs (snoRNAs) is associated with rRNA biogenesis in the nucleolar region (Kass *et al.*, 1990), where the transcription of the rRNA genes occurs.

The excision of certain introns in RNA can also occur as the result of RNA self-cleavage. Catalytic RNAs, generically referred to as ribozymes were first described by Cech *et al.* (1981); extensive information on this fascinating topic in RNA biology is described elsewhere (for reviews, see Doherty and Doudna, 2000; Kurreck, 2003; Blount and Uhlenbeck, 2005). In particular, the group I intron ribozyme and RNase P ribozyme are well-known because they are first two ribozymes to be discovered (Kruger *et al.*, 1982; Guerrier-Takada and Altman, 1984). These discoveries led to the awarding of the Nobel Prize in 1989 to Thomas Cech and Sidney Altman.

The immediate products of transcription are RNA molecules whose sequence correlates precisely with the DNA from which it is derived. These primary transcription products are only a precursor to functional RNA, and are confined to the eukaryotic nucleus. Collectively, then, all precursor RNAs in the nucleus are known as *heterogeneous nuclear RNA* (hnRNA). hnRNA and specific nuclear proteins form heterogeneous nuclear ribonucleoprotein complexes (hnRNPs), an extremely abundant component of the nucleus (for review, see Dreyfuss *et al.*, 2002). For mRNA to form and function as a template for the synthesis of its encoded protein, the introns must first be removed as described above. The resulting concatenation of exon sequences is but one of several modifications to which the precursor RNA is subjected on the road to RNA maturation. This

type of processing, and the steps that follow, occur in the nucleus. Thoroughly processed transcripts are exported to the cytoplasm where they then function as mature mRNA molecules while the intervening intron sequences are reserved in the nucleus and then degraded. Although intron removal and the splicing together of exons in and of itself is not required for transport from the nucleus, since intron-less transcripts move efficiently into the cytoplasm, splicing clearly enhances transport.

RNA polymerases and the products of transcription

Genes are transcribed by enzymes known as RNA polymerases, thereby producing the major types of RNA, including ribosomal RNA (rRNA), transfer RNA (tRNA), and mRNA, as well as all of the smaller RNA species. Eukaryotic genes are transcribed by one of four nuclear RNA polymerases; these enzymes are among the largest and most complex proteins in the cell and consist of more subunits than their prokaryotic counterpart. The eukaryotic enzymes are properly known as RNA polymerases I, II, III, and IV, each of which is responsible for transcribing a different class of genes (Table 1.4). Prokaryotes, in contrast, exhibit only one type of RNA polymerase, which transcribes all classes of RNA. RNA polymerases are active only in the presence of DNA, and require the nucleotides ATP, CTP, GTP, and UTP as precursors, myriad transcription factors, and Mg^{++} . The transcriptional products of eukaryotic RNA polymerases I, II, and III may be distinguished by their differential sensitivity to the bicyclic octapeptide fungal toxin α -amanitin¹⁰ (Marzluff

Table 1.4 Eukaryotic RNA Polymerase Enzymes and Their Respective Products and Sensitivities to α -Amanitin

Eukaryotic enzyme	Products	Sensitivity to α -amanitin
RNA Polymerase I	rRNA (28S, 18S, 5.8S)	–
RNA Polymerase II	pre-mRNA → → → mRNA; miRNA ¹	++ + +
RNA Polymerase III	tRNA, 5S rRNA, snRNA, snoRNA, scRNA, miRNA	+
RNA Polymerase IV		
spRNAP-IV (mammalian)	mRNA (sub-set)	–
Pol IV (plants)	siRNA	–
RdRPs (RNA-dependent RNA polymerases)	Amplification of miRNA for gene silencing	–

¹While most known miRNAs are transcribed by RNA polymerase II, an increasing number of these important regulatory transcripts have been shown to be transcribed by RNA polymerase III.

¹⁰Extracted from the poisonous mushroom *Amanita phalloides*.

and Huang, 1984; Roeder, 1976), applications for which are described in Chapter 16.

As is the case in the synthesis of all nucleic acid molecules, RNA transcripts are assembled only in the 5'→3' direction. Transcription involves three distinct phases, namely, initiation, elongation, and termination, all of which have been described in great detail elsewhere and the details of which are beyond the scope of this volume. Briefly, initiation involves the attachment of RNA polymerase to a DNA template promoter, via transcription-, activation-, and initiation factors, followed by the acquisition of what will be the first ribonucleotide in the RNA molecule. Elongation involves the sequential addition of ribonucleotides to the nascent chain, a process also involving accessory protein elongation factors. Termination is the completion of RNA synthesis, whether appropriately or prematurely, and the disengagement of both the newly synthesized RNA and the RNA polymerase from the DNA template. Transcription termination, as with initiation and elongation, is sequence-dependent and is influenced by the presence of small proteins (termination factors) as well as the transient formation of RNA 2° structures. Mutations notwithstanding, the nucleotide sequence of the resulting RNA molecule is identical to the coding strand of the DNA from which it is derived, the only difference being the substitution of the base uracil for thymine.

In eukaryotic cells, transcription of the genes encoding ribosomal RNA is mediated by RNA polymerase I. The primary product of RNA polymerase I transcription in higher eukaryotes, a large unspliced 45S precursor RNA¹¹, eventually yields the smaller 28S, 18S and 5.8S rRNAs after processing (Fig. 1.6). RNA polymerase III is responsible for transcribing the genes that encode tRNA

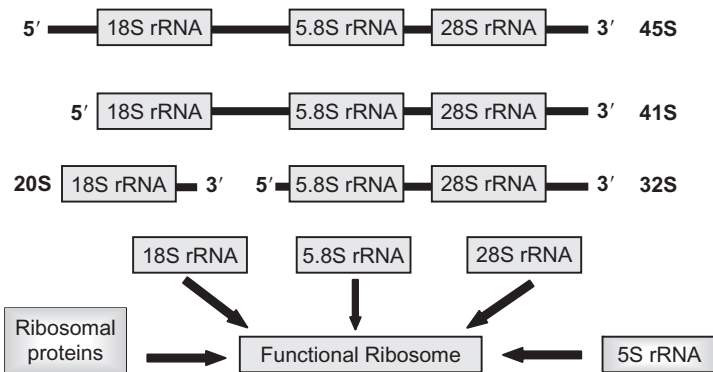


Figure 1.6 rRNA biogenesis in human cells. The 28S, 18S, and 5.8S rRNA are liberated from a common 45S transcript that is the product of RNA polymerase I. The other essential transcript, 5S mRNA, is produced independently by RNA polymerase III. Adopted, in part, from Lewin, *Genes VI* (1997), by permission of Oxford University Press.

¹¹The 45S rRNA is the largest known precursor RNA in mammals.

molecules, the 5S rRNA, certain repetitive elements and the snRNA, snoRNA, and scRNA transcripts described above.

The typical mammalian cell contains approximately 1 to 5×10^{-5} μg RNA, a vast majority of which represents the transcriptional products of RNA polymerase I and RNA polymerase III. Between 80–85% of cellular RNA is found in ribosomes in the form of the 28S, 18S, 5.8S, and 5S rRNAs; these transcripts form complexes with myriad ribosome-specific proteins to form the 60S and 40S eukaryotic ribosomal subunits, respectively. As described in Chapter 9, the abundant 28S and 18S rRNAs serve as excellent natural molecular weight size markers when either total cellular RNA or total cytoplasmic RNA is electrophoresed. A comparison of eukaryotic ribosomes and other aspects of transcription and translation, with their prokaryotic counterparts is presented in Table 1.5. tRNA constitutes 10–15% of the total cellular RNA mass, while all of the mRNAs collectively account for only 1–4% of the total RNA in a cell.

Historically, rRNA was first believed to have a template role in the synthesis of proteins. The first indication that a new, separate class of RNA, mRNA, acted as

Table 1.5 Comparison of Transcription and Translation in Prokaryotic and Eukaryotic Cells

		Prokaryotes	Eukaryotes
Transcription	Gene organization	Mostly polycistronic (operon model); some are monocistronic	Nearly all are monocistronic; polycistronic and bicistronic mRNAs observed in plants
	Location	Single cellular compartment	Nucleus
	mRNA half-life	2 min; range 1–10 min	10 h; range 1–24 h
	Transcription rate	40 nts s^{-1}	40 nts s^{-1}
	Posttranscriptional modifications	Limited-to-none	Extensive, including 5' capping, intron removal, and polyadenylation
	mRNA average size	0.9–1.5 kb	1.9–2.2 kb
Translation	Intact ribosomes	70S	80S
	Ribosomal subunits	50S, 30S	60S, 40S
	rRNA species	23S, 16S, 5S	28S, 18S, 5.8S, 5S
	Location	Single cellular compartment	Cytoplasm
	Translation rate	15 amino acids per sec	5–8 amino acids per sec
	Posttranslational modifications	Limited-to-none	Extensive

This table does not include transcription and translation information associated with mitochondrial or chloroplast genes.

the template molecule emerged from studies involving T4 phage-infected *E. coli* cells (Volkin and Astrachan, 1956; Brenner *et al.*, 1961; Gros *et al.*, 1961; Hall and Spiegelman, 1961). Later, multiple 5' ends were observed for the late SV40 mRNAs (Ghosh *et al.*, 1978) and late polyoma mRNAs (Flavell *et al.*, 1979), which were the first indications of the heterogeneous nature of RNA polymerase II transcription initiation.

The transcription products of RNA polymerases I and III are rather undiversified and generally not subject to change as a function of the cell state. The transcription products of RNA polymerase II, however, are as diverse as the cellular biochemistry itself. This is not at all unexpected because the mRNA in great measure drives the phenotype of the cell. The relative contribution of RNA polymerase II transcription products is generally between 20 and 40% depending on cell type and cell state, but only 1–4% of which is mature mRNA. The genes that, upon transcription and processing, render functional mRNA are localized mostly within the non-repetitive regions of the genome, as demonstrated by R₀t kinetics studies. Whereas a major portion of the genome¹² is believed to be transcriptionally silent, a typical cell is transcribing several thousand different genes at any given moment. Given the complexity of the cellular biochemistry, this observed heterogeneity within the mRNA population is necessary to satisfy even basal-level requirements for viability. Interestingly, as much as 70% of all transcribed hnRNA is degraded in the nucleus, providing a rationale for studying posttranscriptional regulation of gene expression.

In mammals, a fourth RNA polymerase (single-polypeptide nuclear RNA polymerase IV [spRNAP-IV]) has recently been described (Kravchenko *et al.*, 2005). This polypeptide is encoded by a nuclear gene that ordinarily encodes a mitochondrial RNA polymerase. As a consequence of alternative splicing, a truncated polypeptide results which lacks 262 amino acids near the amino terminus, including the mitochondrial transit sequence. Thus, spRNAP-IV is located in the nucleus and regulates a subset of more than one thousand nuclear-encoded genes, though the level of cooperation between RNA polymerase II and spRNAP-IV in gene transcription remains unclear. Suppression of the spRNAP-IV gene seriously inhibits cell growth, gradually leading the cell toward the apoptotic pathway.

Similarly, a novel polymerase, RNA polymerase IV (RNAP IV, or simply RNA pol IV), has been discovered in plants (Onodera *et al.*, 2005). Unlike spRNAP-IV in animals, Pol IV does not appear to be necessary for viability, and it exists in two distinct forms. RNAP IV does not appear to overlap with the known functions of RNA polymerases I, II, and III; in contrast, it is used by the cell to promote methylation-associated higher-order heterochromatin formation (Herr *et al.*, 2005; Kanno *et al.*, 2005). Recent data have demonstrated that, in *Arabidopsis*, RNAP IV is required for the synthesis of an overwhelming majority of all siRNAs which, in turn, direct RNA-mediated DNA methylation events that are ultimately associated with gene silencing (Zhang *et al.*, 2007).

¹²Diploid genome size is 6.6×10^9 bp in *Homo sapiens*.

Another type of RNA polymerase, RNA-dependent RNA polymerase (RdRP), also known as RNA replicase, is an enzyme that can synthesize RNA from an RNA template, rather than the requisite DNA template associated with the RNA polymerase I, II, III, and IV. RdRP is well known for its role in poliovirus replication, and has recently come under scrutiny because of the documented role of this enzyme in the RNA interference pathways, especially in higher plants. This aspect of RdRP function is discussed in detail in Chapter 23 (RNA Interference).

Messenger RNA

A great many genes are transcribed constitutively by RNA polymerase II¹³, and it is clear that large quantities of heterogeneous nuclear RNA (hnRNA) are turned over in the nucleus. In eukaryotic cells, messenger RNAs (mRNA) are derived from precursor hnRNA through a series of modifying reactions, which include formation of the 5' cap, methylation, splicing, 3'-end processing, and frequently, polyadenylation. Only 1–3% of the total RNA in the cytoplasm of a typical eukaryotic cell is mature mRNA. RNA is produced at different rates from different loci; therefore, each mRNA species is classified based on its cytoplasmic prevalence or, more properly, its abundance. There are three official such categories, high abundance, medium abundance, and low abundance mRNAs and, in the mind of this Author, the unofficial *very low* abundance category.

Highly abundant transcripts are present in hundreds of copies per cell. These are most often observed when a cell is producing an enormous quantity of a particular protein or is high specialized or *differentiated* to perform a unique function. Medium abundance transcripts are best thought of as being present in dozens of copies per cell; many genes with housekeeping¹⁴ functions manifest their mRNAs at this level of prevalence in the cell. Low abundance mRNAs are generally prevalent in 10 or fewer copies per cell and often are difficult to assay by many of the older classical techniques, such as Northern analysis (Chapter 11), without some form of enrichment in order to increase the statistical probability that such rare messages will be detectable. *Very low* abundance mRNAs are those present in fewer than one copy per cell, a designation which generally is generally associated with heterogeneous tissue samples or, very commonly, in cases where cancer cells growing in culture manifest a variable, heterogeneous karyotype. In the past these types of mRNAs were referred to as the “hard to clone genes”, though newer methods for the assay of gene expression have

¹³RNA polymerase II is also involved in orchestrating splicing and polyadenylation of hnRNA (reviewed by Hirose and Manley, 2000).

¹⁴A gene is said to have a housekeeping function if the encoded protein plays a maintenance role or if basal expression is needed to maintain the physiology of the cell or even viability. Since the expression of these genes is generally not expected to change as a function of cell state, housekeeping genes are assayed as internal controls in various quantitative assays. See Chapter 19 for further details.

revealed a plethora of previously unknown transcripts of various persuasions. Most importantly, it is essential to realize that the prevalence or abundance of an mRNA species in a cell is subject to change of monumental proportions. Such changes may occur in response to natural changes in the cellular milieu or due to experimental manipulation.

Topology of a typical mRNA molecule

A typical human fibroblast cell contains approximately 1 picogram (pg) of mRNA, which is equivalent to about 10^6 molecules transcribed from a particular subset of an estimated 25,000–30,000 genes. While this mRNA heterogeneity reflects the diversity of proteins that these mRNAs encode, a typical eukaryotic mRNA molecule shares several topological features with nearly all other mRNA molecules (Fig. 1.7). As will become evident from the descriptions which follow, producing a function mRNA molecule is amazingly complex.

5' Cap

The overwhelming majority of mature eukaryotic mRNA molecules are characteristically monocistronic polyribonucleotides produced by RNA polymerase II. The immediate product of transcription, a precursor hnRNA molecule, displays the following structure at the 5' end of the molecule



meaning that the first *transcribed* nucleotide is usually a purine, either adenosine or guanosine, and the 5' triphosphate of which remains intact. Phosphodiester bonds join adjacent ribonucleotides 5'-3'. A structure known as the 5' cap (Reddy *et al.*, 1974; Shatkin, 1976; Banerjee, 1980) is then assembled in a step-wise manner just after the initiation of transcription. Capping occurs as soon as the mRNA emerges from the RNA polymerase complex, when nascent polynucleotides are generally between 20–30 bases long

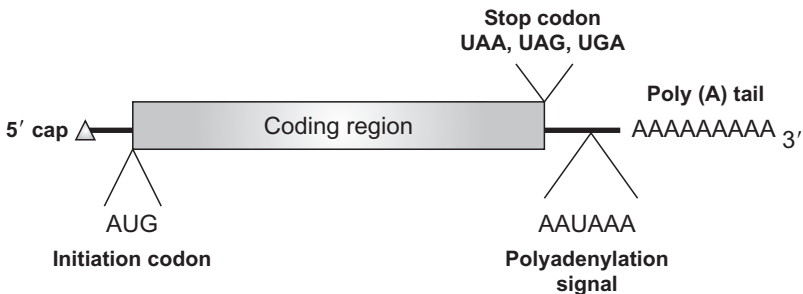


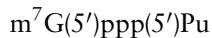
Figure 1.7 Topology of a typical eukaryotic mRNA molecule.

(Coppola *et al.*, 1983). The cap consists of a rare 5'-5' triphosphate linkage between the first transcribed nucleotide and a 7-methylguanosine (m^7G) nucleotide, the polarity of which effectively seals the 5' end of the transcript.

Addition of the terminal guanosine nucleotide (G) to the 5' end of the nascent transcript occurs in the nucleus, and first requires removal of the γ -phosphate from the 5' end of the hnRNA. Subsequently, the terminal "G" is added by the enzyme guanylyltransferase and results in the structure



in which the new terminal guanosine nucleotide is joined 5'-5' to what was the first transcribed nucleotide via what is best thought of as an inverted linkage. The resulting 5' cap structure is then subjected to one or more methylation events¹⁵, the first of which is directed toward the number 7 position of the terminal guanine, courtesy of the enzyme guanine-7-methyltransferase, and is then represented as



The majority of eukaryotic mRNAs subsequently experience an additional methylation directed toward the 2' oxygen in the ribose of the second nucleotide, catalyzed by the enzyme 2'-O-methyl-transferase. Depending on the mRNA, as many as four methylation events can occur involving the 5' guanine nucleotide as well as the base and/or sugar of the penultimate- and subpenultimate nucleotides and may include the formation of N⁶-methyladenosine. Capping of eukaryotic mRNAs also precedes rare internal mRNA adenylate methylation; when it does occur this infrequent modification results in the formation of N⁶-methyladenosine residues early in mRNA biogenesis and which are conserved during RNA processing (Shatkin, 1976; Chen-Kiang *et al.*, 1979). The extent of methylation is a function of mRNA species, and higher organisms usually have more extensively methylated caps. Cap structures with multiple methyl groups have also been observed in the small nuclear RNA species (Furuichi and Shatkin, 1989).

5'-capping is only one example of a posttranscriptional modification associated with cellular gene expression and is observed in virtually all eukaryotic mRNAs. The presence of the 5' cap is required to support initiation of translation in eukaryotes; mRNAs produced by *in vitro* transcription must be subjected to an *in vitro* capping reaction if the transcripts are to be expected to support synthesis of the encoded protein. Formation of the translation apparatus is initiated in part by cap-binding proteins (for review, see Lewin, 2008), followed by the assembly of the ribosomal subunits as mediated by initiation

¹⁵The donor of the methyl groups in the cap is S-adenosyl-methionine (SAM).

factors. In short, this is how ribosomes recognize those transcripts that are to be translated: rRNA and tRNA are non-capped.

Capping apparently also confers transcript stability by protecting against phosphatase attack and 5'→3' exonucleolytic degradation (Furuichi *et al.*, 1977; Furuichi and Shatkin, 1989). In contrast, prokaryotic mRNAs, which naturally lack a 5' cap structure, are degraded exonucleolytically from the 5' end even while translation is ongoing downstream. Nor is the 5' cap a characteristic of mitochondrial or chloroplast mRNAs. Most animal viruses that replicate in eukaryotic cells manifest 5' capped mRNAs, a noteworthy exception being poliovirus (Hewlett *et al.*, 1976; Nomoto *et al.*, 1976).

Leader sequence

The first nucleotides immediately 3' to the eukaryotic cap structure constitute the non-translated leader sequence or, more formally, the 5' untranslated region (UTR). In eukaryotic systems, the typical leader sequence is 30–50 bases long, though leader sequences as long as hundreds of bases long or as short as three bases have been observed. As its name implies, this variable-length region is not translated. According to the ribosome scanning model (Kozak, 1978, 1989), the 40S subunit binds to the 5' cap and then travels along the leader sequence, pausing only when reaching the first AUG codon, which usually demarcates the boundary between the leader sequence and the coding region. The initial association between the 5' cap and the 40S subunit is mediated by cap binding proteins while the eventual assembly of the functional ribosome is mediated by eukaryotic initiation factors (eIFs) and other mRNA binding proteins (Leibold and Munro, 1988; Dever, 2002).

Coding region

The structural portion of the mRNA, that is, the coding region, begins with an initiation codon, most often the triplet AUG. Far less frequently, translation in eukaryotic cells may be initiated from non-AUG start codons (Hann *et al.*, 1988; Florkiewicz and Sommer, 1989), including GUG, ACG, and CUG. Curiously, these non-canonical initiation codons differ from AUG by only one base. Prokaryotes, in contrast, seem to like both GUG and AUG for the purpose of translation initiation. In addition to this highly conserved three-base motif, a putative initiation codon must be read in “good context”. Early on it was thought that a purine, usually adenine, present at the position three nucleotides upstream from the AUG codon was all that was required in order for the initiation codon to be recognized by the ribosome (Kozak, 1986). It is now clear that a purine three bases before AUG and a G immediately thereafter is necessary for optimal initiation of translation; these flanking structural elements are presumed to be indicators pointing to the initiation codon.

An AUG codon at the beginning of the coding region results in the placement of the amino acid methionine at the amino terminus of eukaryotic cytoplasmic proteins, and the placement of formyl-methionine at the amino terminus of

proteins in prokaryotic cells and in eukaryotic organelles. Every three nucleotides beyond the initiation codon specifies the placement of another amino acid into the nascent polypeptide. It has been suggested that eukaryotic peptide elongation may also stall along specific stretches of the mRNA template (Woolin and Walter, 1988), enhancing polysome formation and thereby producing a greater number of polypeptides from a single mRNA molecule. Elongation continues in this manner until a nonsense (stop) codon (UAA, UAG, or UGA) is encountered which, in the absence of tRNA suppressor mutations, is the normal signal to halt translation. Thus, the 5' end of the mRNA coding region corresponds to the amino terminus, and the 3' end corresponds to the carboxy terminus of the primary polypeptide that results from translation of the mRNA. Typical eukaryotic mRNA coding regions range from 200–1500 nucleotides long, though extremely short and extremely long mRNA coding regions have been identified.

Trailer sequence

Beyond the stop codon lies another non-translated area known as the trailer sequence or simply the 3' UTR. This sequence usually ranges from 50–150 nucleotides long, although trailer sequences as long as 1000 bases have been reported. Further, different mRNA molecules encoding the same polypeptide may differ only with respect to the length of their 3' trailer region, of which dihydrofolate reductase mRNA is one example (Will and Dolnick, 1989; Kubo *et al.*, 2006). Although no function has yet been ascribed to the trailer region, it is believed to influence the stability of the molecule in the cytoplasm due to *cis*-acting elements which bind various proteins and that may also influence mRNA localization. It is worth noting that most trailer regions manifest a highly conserved AAUAAA sequence within 30 nucleotides of the location where the synthesis of the poly(A) tail will begin (Proudfoot and Brownlee, 1976). Newer evidence suggests that microRNA (miRNA) molecules may target the trailer sequence and, in so doing, influence gene expression (Lai, 2002).

Poly(A) tail

Most eukaryotic mRNA species are further characterized by a 50–250 nucleotide tract of polyadenylic acid at the 3' end of the molecule (Lim and Cannelakis, 1970; Edmonds *et al.*, 1971; Brawerman, 1976). This structure is known as the poly(A) tail, the precise length of which is a function of mRNA species and translational status in the cell (Kuge and Richter, 1995; reviewed by Wahle and Ruegsegger, 1999; Hunt, 2007). The poly(A) structure is added enzymatically to precursor hnRNA very soon after transcription, by one of a family of nuclear enzymes known as polyadenylate polymerases, or simply poly(A) polymerase, and as many as fifteen associated subunits. All poly(A) polymerases show tight specificity for ATP and require a polyribonucleotide with a free 3'-hydroxyl group to which AMP residues are attached sequentially.

There is no poly(T) tract found on the template strand of the associated gene: the poly(A) tail is added in a non-template dependent manner.

The mRNA fraction that exhibits this feature is collectively described as poly(A)⁺ mRNA, and nearly all (>99%) types of eukaryotic mRNAs are believed to exhibit the poly(A) structure. When polyadenylation is inhibited¹⁶, then normally adenylated mRNAs fail to emerge from the hnRNA pool. Interestingly, fully one-third of the total *mass* of cellular mRNA lacks a 3' poly(A) tail, of which the overwhelming component are the histone mRNAs, which encode major chromosomal structural proteins. If the poly(A) tail is needed for nucleocytoplasmic transport, then non-adenylated messages must use an alternative method for nuclear egress, presumably based on features that are unique to mature histone transcripts. It is clear that, regardless of the structure of the 3' end of an mRNA, movement through the pores in the nuclear envelop occurs when specific proteins bind to the mRNA and, in so doing, form a messenger ribonucleoprotein (mRNP) particle. It is widely believed that improperly spliced or otherwise compromised mRNAs fail to associate with the correct combination of proteins required for nucleocytoplasmic transport, thereby promoting their retention in the nucleus. Once in the cytoplasm an mRNA may, or may not, be translated, depending on local conditions.

The modulation of polyadenylation of many mRNAs has been observed and a rapidly growing body of evidence supports a role for the poly(A) tail both in the stability of a particular species in the cytoplasm as well as its translational efficiency. It is likely that the specialized 5' and 3' features of eukaryotic mRNA may protect these molecules, at least in part, from exonuclease digestion. On the other hand, sequence information near the 5' and/or near the 3' end may facilitate degradation of these molecules in response to various cytoplasmic cues (Jackson and Standart, 1990; Spirin, 1994; Standart and Jackson, 1994; Curtis *et al.*, 1995).

Transcription apparently does not terminate at the site where polyadenylation begins, but in some cases as much as 1000 bp or more downstream (that is, in the 3' direction). This phenomenon was first observed in the adenovirus model system (Nevins and Darnell, 1978; Fraser *et al.*, 1979). The actual site of polyadenylation is generated by stringently regulated splicing within the trailer region of the primary transcript, 3' to which the poly(A) tail is added. Polyadenylation efficiency is heavily dependent on the highly conserved AAUAAA motif, known more commonly as the poly(A) signal, another somewhat conserved GU-rich sequence located downstream from the polyadenylation start site (upstream from the polyadenylation site in plants), and other sequences upstream of the AAUAAA motif. These features, or subtle variants, are widely observable among the eukaryotes, with the exception of yeast. The precise spatial relationship of these sequences and degree of sequence conservation direct the cleavage of an hnRNA molecule downstream from the AAUAAA signal, and subsequent poly(A) polymerization to the 3'-hydroxyl

¹⁶Cordycepin (3'-deoxyadenosine) prevents natural polyadenylation from occurring (Zeevi *et al.*, 1982) but does not interfere with transcription in general or with the synthesis of hnRNA.

group that is generated by that cleavage event. These posttranscriptional modifications have been shown to be mediated, at least in part, by small nuclear ribonucleoproteins (snRNPs) (Birnstiel *et al.*, 1985; Berget and Robberson, 1986; Black and Steitz, 1986; Hashimoto and Steitz, 1986).

The importance of the polyadenylation signal is easily demonstrated (Fitzgerald and Shenk, 1981; Manley *et al.*, 1985): natural- and site-directed mutations and spatial rearrangements interfere profoundly with the formation of poly(A)⁺ mRNA. The polyadenylation signal, while highly conserved, is not a universal signal. For example, the polyadenylation signal AUUAAA has been observed in chicken lysozyme mRNA (Jung *et al.*, 1980) and in a one of several adenovirus mRNAs (Ahmed *et al.*, 1982); the sequence AAUUA has likewise been found in a minor form of pancreatic α -amylase mRNA (Tosi *et al.*, 1981). Further, a motif for short-lived mRNAs, AUUUA, has been observed in certain oncogene- and cytokine transcripts (Shaw and Kamen, 1986). Such a sequence, now known as an ARE (AU-rich element), is commonly observed in unstable mRNAs, and is believed to facilitate the degradation of mRNAs so-endowed. This pentanucleotide sequence, localized in the non-translated trailer region, may be repeated in tandem several times. ARE-mediated destabilization occurs by deadenylation, i.e., cleavage of the poly(A) tail, followed by endonuclease-mediated degradation. It now appears that polyadenylation may be involved in the regulation mRNA stability in prokaryotes (O'Hara *et al.*, 1995), too.

The highly conserved nature of the polyadenylation signal and the abundance of adenylated eukaryotic transcripts suggests a central regulatory role in gene expression. Polyadenylate-binding protein (PABP)¹⁷ appears to bind to every 15–20 bases within the poly(A) tail, the length of which correlates well with the stability of the mRNA as well its translational efficiency. This is mediated in part by the formation a transient closed loop complex in which the 5' end and the 3' end of the same mRNA molecule are brought into proximity by eukaryotic translation initiation factor 4G (eIF4G) and which involves other translation-associated factors as well.

Organellar mRNAs

It has been recognized for some time that both mitochondria and chloroplasts have their own chromosome(s), circular in form, which are inherited independently of nuclear chromatin and in a non-Mendelian fashion. The genomes of these organelles are referred to as mtDNA and ctDNA, respectively. The genes found in mtDNA and ctDNA encode proteins that are uniquely localized in these organelles, though mitochondria and chloroplasts each import proteins encoded by nuclear genes to support normal organellar physiology.

Although the mechanics of transcription and translation are highly conserved, certain structural differences distinguish mRNA species confined to organelles.

¹⁷Methodologies for characterization of polyadenylate-binding proteins are described by Sachs and Kornberg (1990) and Görlach *et al.* (1994).

For example, neither mitochondrial nor chloroplast mRNAs exhibit a 5' cap structure. Most mitochondrial transcripts are constitutively polyadenylated at the 3' terminus, while most chloroplast mRNAs are not polyadenylated. In the case of both organelles, the poly(A) tail, when present, is usually quite short by comparison with adenylated cytoplasmic transcripts. Curiously, the effect of polyadenylation on transcripts in the chloroplast is destabilizing (Lisitsky *et al.*, 1997; Gagliardi and Leaver, 1999; reviewed by Hayes *et al.*, 1999; Schuster *et al.*, 1999), rather than stabilizing as seen in cytoplasmic mRNAs (Manley and Proudfoot, 1994). Mitochondrial mRNAs frequently utilize the non-canonical AUA and AUU translation start codons, and these are usually observed very close to the 5' terminus. Interestingly, the mRNAs found in the mitochondria of yeast are more closely related to typical cytoplasmic mRNAs, structurally speaking, than to mitochondrial mRNAs observed in other cell types.

mRNA stability, transport, and turnover

A fundamental regulator of gene expression in all cell types and in all subcellular compartments is the stability of translatable transcripts. In general, mRNAs do not have long half-lives (Table 1.5), presumably so as to prevent the overproduction of a normal protein which could, in turn, disrupt homeostasis and give rise to a disease state. At the same time, mRNAs must remain stable long enough to become recognized and engaged by the translation apparatus, which are the intrinsic functions of the 5' cap. If large quantities of a protein are to be produced in a cell, one may expect that the corresponding gene will be transcribed with a greater frequency than other genes with, for example, housekeeping functions. Similarly, the formation of mRNA secondary structures close to the 5' end in both plants and animals can severely limit the scanning of the mRNA such that the initiation of translation is all but inhibited (Pain, 1996; Kozak, 1991; Dinesh-Kumar and Miller, 1993; Futterer and Hohn, 1996; for review, see Kozak, 1999).

At the other end of the molecule, the length of the poly(A) tail itself plays a role in mRNA stability, as shortening of the poly(A) tail results in destabilization of cytoplasmic transcripts (Decker and Parker, 1994; Beelman and Parker, 1995). Early studies demonstrated that the enzymatic removal of the poly(A) tract from globin mRNA results in a rapid loss of translatability in frog oocytes due to rapid degradation (Huez *et al.*, 1974; Marbaix *et al.*, 1975). More recent studies have demonstrated the role of the 3' AREs; deletion of these sequences greatly reduces the rate of deadenylation, thereby prolonging mRNA in the cytoplasm (Wilson and Treisman, 1988; Shyu *et al.*, 1991; Decker and Parker, 1993; Chen and Shyu, 1994). Further, an increasing body of evidence is suggesting that both the length and nucleotide composition of the 5' UTR and 3' UTR play a previously unrecognized role in the stability of the transcript (reviewed by Lewin, 2008). Finally, another recently discovered pathway known as nonsense-mediated mRNA decay appears to be at work in eukaryotic cells which rapidly targets for degradation mRNAs that have acquired by mutation

an additional stop (nonsense) codon upstream of the naturally occurring stop codon (Le Hir *et al.*, 2000; Lykke-Andersen *et al.*, 2000).

RNA molecules in the eukaryotic nucleus form ribonucleoprotein particles cotranscriptionally to facilitate nuclear egress, which is believed to occur in a 5' to 3' manner. Some RNA molecules are transported into mitochondria via the cytoplasm, including some tRNAs needed to support the synthesis of mitochondrial proteins. Thus, one may expect to isolate different populations of transcripts depending on the method of cell lysis and concomitant extent of organellar disruption, as described in Chapter 2.

mRNA in plants is also known to travel from one cell to another via the phloem, the first demonstration of which was in tomato. Briefly, wild type tomato plants were grafted onto plants exhibiting the dominant *Me* phenotype, a mutation that affects leaf morphology. mRNA from the mutant rootstock was shown to travel to the grafted wild type plants and alter their appearance of the leaves (Kim *et al.*, 2001). Similarly, intercellular mRNA transport is well established in animal models, especially during embryogenesis in *Drosophila*.

Bicistronic mRNAs

Although the one mRNA, one polypeptide relationship is widespread among eukaryotes, bicistronic mRNAs have been identified in certain organisms. A bicistronic mRNA is capable of directing the synthesis of two different proteins. One might think of bicistronic mRNAs as the eukaryotic answer to the polycistronic mRNAs that are nearly universally observed among prokaryotes. Taking this a step further, functional tricistronic mRNAs, encoding three different polypeptides, are in use in certain *in vitro* applications. Due to the peculiarities associated with translation in eukaryotes, the first (upstream) encoded protein is synthesized in the 5'-cap dependent manner usually associated with translation of monocistronic mRNAs while initiation of the synthesis of the second (downstream) polypeptide is under the control by an internal ribosome entry site (IRES) which allows ribosome assembly in a non-cap-dependent manner. This translation strategy is widespread among eukaryotic viruses and, while still considered a rarity in higher animal cells, there are reports of bicistronic mRNAs in plants, including the tomato *tomPro1* locus and in *Arabidopsis* (see Farrell and Bassett, 2007 for a recent review). It is also possible for a single-reading-frame mRNA to produce two or more polypeptides by cleavage of large precursor protein (a zymogen, for example), such as with the animal hormones oxytocin and vasopressin (Richter, 1983). These observations have lead investigators to rethink the entire process of the regulation of gene expression and to analyze gene expression data circumspectly.

Prokaryotic mRNAs

The mechanics and goals of transcription and translation in the prokaryotic cell are similar to these processes in eukaryotes, namely the manifestation of gene

expression. There are, however, structural differences that are readily observable when comparing transcripts from these two cell types. As described above, prokaryotic mRNAs are polycistronic, encoding more than one polypeptide. These transcripts often experience polysome-associated translation and simultaneous 5' degradation even prior to the completion of transcription at the 3' end. While endoribonucleases are at work destroying RNA from within the molecule, a process initiated by RNase E, exoribonucleases are also at work disassembling the RNase E cleavage products one base at a time from the free 3' end.

Remarkably, polyadenylation may play a role in mRNA in degradation in bacteria, in very sharp contrast to the understood role of this mRNA feature in eukaryotes. It appears that the prokaryotic version of poly(A) polymerase adds as many as 40 adenine nucleotides to some mRNAs as if in some way marking them for obliteration.

More recently, riboswitches have attracted a great deal of attention as modulators of gene expression in prokaryotes. A riboswitch is a double-stranded region of an RNA molecule, formed by intramolecular base-pairing, that is able to bind small metabolites which, in turn, directly regulates the activity of the gene from which the RNA was originally transcribed (Nahvi, *et al.*, 2002; Winkler, Cohen, and Breaker, 2002; Winkler *et al.*, 2002; Mandal and Breaker, 2004; reviewed by Montange and Batey, 2008). Succinctly, an mRNA riboswitch regulates its own expression. Although most riboswitches are prokaryotic, at least one type of riboswitch has been found in plants and fungi.

mRNA sequence and structure affect translation

Translation is a series of variegated biochemical events acting concertedly to sustain the production of proteins. The initiation, elongation, and termination of this process are sequence-dependent (e.g. the initiation and nonsense codons), and mutations therein can have strong negative consequences on the final manifestation of gene expression. Further, IRESs can initiate translation by promoting ribosome binding within a transcript and initiate translation without a start codon. It is important to note, however, that while features of the transcript sequence are certainly intimately linked to functionality, the three-dimensional shape(s) that a transcript can form also influence translatability profoundly, both *in vivo* and *in vitro*.

Translation results when an engaged ribosome is unencumbered as it moves along a transcript until it reaches a stop codon. The formation of secondary structure(s) involving the 5' leader sequence of the mRNA can limit or altogether suppress translation and this mode of gene regulation has been observed in both plant and animal models, presumably to prevent the over-expression of a specific protein. Moderately stable hairpins upstream of an AUG codon and close to the 5' end of a transcript are known to suppress translation efficiency, presumably by impacting the ability of ribosomes to load onto the mRNA. Mutation and deletion experiments that reduce hairpin stability generally result in an increase in the amount of the encoded protein. Interestingly, a -19 kcal/mol

hairpin 14 nts downstream from an AUG codon can actually enhance translation initiation (Kozak, 1990), possibly because the hairpin causes strategic pausing of the ribosome directly over the start codon, thereby enhancing translation initiation.

Recent reports suggest that the 3' UTR (the trailer sequence of an mRNA), may have a profound role regulating translation initiation. Because eukaryotic mRNAs acquire a circular conformation, bringing the 3' end of the transcript close to the 5' cap and the 5'-UTR, it appears that either the association of proteins, the formation of secondary structures within the 3'-UTR, or perhaps both, influence overall translation efficiency. This translation regulatory mechanism appears to be widespread in animals (Allard *et al.*, 2005; Wax *et al.*, 2005), in plants (Browning, 1996), and among plant viruses (Guo *et al.*, 2001).

Levels of gene regulation

Controlling gene expression requires interfering with either transcription, translation, or the biogenesis of the products of these biochemical processes. The number of potential points of gene regulation in living cells is virtually infinite. In the *broadest* sense, these myriad potential regulatory points may be classified under one of four main headings (Fig. 1.8):

1. Regulation at the transcriptional level.
2. Regulation at the posttranscriptional level.
3. Regulation at the translational level.
4. Regulation at the posttranslational level.

At the subcellular level, eukaryotic RNA molecules are produced by transcription. This is the primary regulation level in the cell: either a gene is transcribed or it is not. Subsequent processing of the primary transcript is a continuation of what is known as mRNA biogenesis. Mature RNA molecules, which have been appropriately modified, and not the intron sequences, which have been removed by excision, may be exported from the nucleus into the cytoplasm. Details as to the nature of this extremely discriminating nucleocytoplasmic transport mechanism, involving movement to and then through nuclear pore complexes, have been described (Piñol-Roma and Dreyfuss, 1993; Maquat, 1997). Once in the cytoplasm, mature messenger RNA may or may not become associated with the protein translational apparatus. It is important to recall here that although a majority of the mRNA is engaged in polysome complexes, the formation of a non-translatable mRNA:protein complex in the cytoplasm is but one possible translation regulation control point. The term "informosome" was used in the past to describe such non-translatable cytoplasmic mRNA:protein complexes (Preobrazhensky and Spirin, 1978; Bag, 1991). If an mRNA molecule does undergo translation, then the primary product of translation, a polypeptide, is usually subjected to extensive posttranslational modifications. Perhaps the best characterized of these is cleavage of the peptide

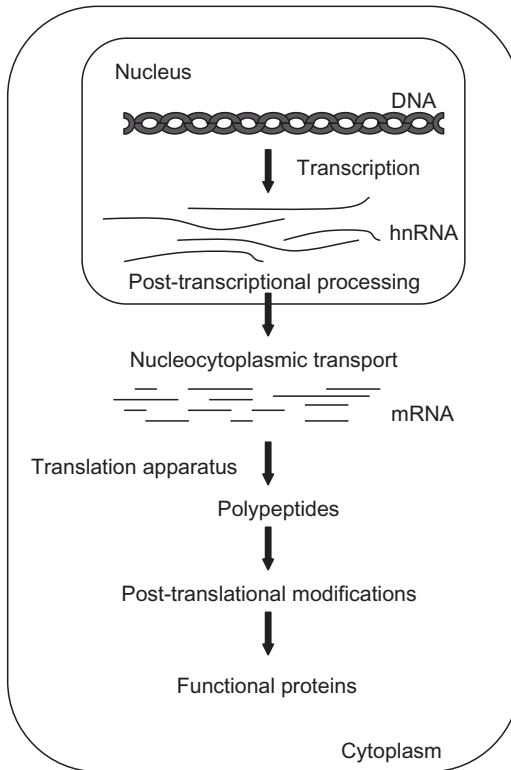


Figure 1.8 Gene expression is subject to regulation at key control points within the cell, beginning with the information encoded in the genome of the cell, through mRNA biogenesis and resulting in final manifestation as a functional protein. See text for details.

signal sequence. Other common posttranslational modifications include, but are not limited to, proteolytic cleavage, acylation, methylation, prenylation, carboxylation, glycosylation, phosphorylation, acetylation, and hydroxylation.

Any variable that influences the efficiency of transcription or prevents transcription from occurring altogether is said to act at the transcriptional level (transcriptional regulation). This may include epigenetic events, such as direct modification of chromatin by DNA methylation or by covalent modification of histone proteins (methylation of lysine 9 in histone 3) that facilitates RNA polymerase accessibility to gene promoter elements resident within the nucleosome by changing the actual shape of the chromatin. Following transcription, any event that influences the splicing of hnRNA (precursor to mRNA), hnRNA stability in the nucleus, nucleocytoplasmic transport, or the very stability of mRNA itself in the cytoplasm is said to act posttranscriptionally (posttranscriptional regulation) and these events are indeed multifarious. For example, while as much as one-half of the mRNA in mammalian cells in tissue culture

may have a half-life of several hours many mRNA species have half-lives of less than one hour, the brevity of which is a natural modulator of gene expression. It is clear that environmental stimuli, too, can influence profoundly the stability of a particular mRNA species (Buckingham *et al.*, 1974; Baumbach *et al.*, 1984; Knight *et al.*, 1985). Stating that gene regulation occurs at a translational or posttranslational level usually implies that a mature mRNA in the cytoplasm is somehow experiencing translational interference at the level of protein synthesis or beyond.

The biochemical profile of a population of cells or a tissue sample is often qualitatively and quantitatively defined as a function of mRNA complexity at the time of cellular disruption. For example, an investigator may wish to determine whether the observed modulation of gene expression in a model system is regulated transcriptionally or by a posttranscriptional event(s). In such a case, the method of cellular disruption and subsequent RNA purification must permit analysis of the nuclear abundance of salient gene transcripts independently of the cytoplasmic abundance of the same. In addition, transcription rate studies require the isolation of intact nuclei that are able to support elongation of, and label incorporation into, initiated transcripts (see nuclear run-on assay, Chapter 16). Therefore, the conditions under which the RNA is purified from biological sources must support the chemical stability of the RNA and ensure RNase inactivation throughout the procedure. Scrupulous, if not compulsive, attention to these details is often the deciding factor in the success or failure of an RNA isolation procedure and subsequent analysis.

Alternative splicing of mRNA from a single genetic locus

Alternative splicing refers to variation in the way that transcripts from a single genetic locus are processed posttranscriptionally. This phenomenon, now widely known in eukaryotic cells, was virtually unheard of 20 years ago. Understanding that higher plant and animal cells can differentially process a particular species of pre-mRNA molecules goes a long way toward explaining how a cell with approximately 20,000 different genes can encode significantly greater numbers of transcripts which, in turn, may be able to encode several hundred thousand (or more) proteins. By joining together various combinations of exons and even changing the functionality of introns, a remarkable diversity of proteins may result from a seeming economy of genomic sequences. Upon completion of the human genome project, the apparent paucity of genes came as a great surprise, given the complexity of the proteome. In the context of alternative transcript processing and posttranslational modification of proteins, the known number of genes in most species is not inconsistent with the biochemical complexity of the cell.

The more common strategies used by cells to accomplish alternative splicing are known as exon skipping, intron retention, the manifestation of cryptic introns, and nonsense-mediated mRNA decay. These phenomena can occur

within a single cell in response to the environmental stimuli or may occur in a tissue-specific manner in order to support the physiology of the organism. Details pertaining to these fascinating posttranscriptional phenomena were recently reviewed (Louzada, 2007). It is important to understand that alternative processing of transcripts may also involve modulating the addition of the poly(A) tail associated with the transcripts' 3' end through the use of alternative polyadenylation sites. Such poly(A) variants may well influence the stability of the transcript in the cell as well as the precise combination of exons which are manifested in the mature mRNA.

Trans-splicing: mRNA repair

An overwhelming majority of higher plant and animal (eukaryotic) genes consist of coding regions known as exons that are separated by intervening, non-coding regions known as introns. In the course of gene expression, transcription results in the synthesis of a large, immature pre-mRNA molecules, consisting of both coding (exon) and non-coding (intron) sequences from a specific gene locus. The process of mRNA maturation involves the removal of introns and the joining together (ligation) of exons so that all of the coding information is contiguous. Generically known as splicing, these well-orchestrated events involve the formation of a spliceosome, i.e. an RNA splicing complex, and involves the removal of introns and the ligation of exons from the same RNA molecule.

Nearly all of the splicing that occurs in the cell, as described above, is known as *cis*-splicing because the exons from a single pre-mRNA molecule are ligated together. In contrast, *trans*-splicing involves the joining of exons from two different RNA molecules, resulting in the formation of a hybrid (chimeric) RNA molecule. Like *cis*-splicing, *trans*-splicing is a naturally occurring process in eukaryotes, albeit at a much, much lower frequency, though it has been reported that as many as 70% of all mRNAs in the nematode *C. elegans* may be subject to *trans*-splicing (reviewed by Hastings, 2005).

Trans-splicing has the potential to be adapted both *in vitro* and *in vivo* to produce an astonishing array of designer proteins, not to mention potential to repair defective mRNAs. SMaRT (spliceosome-mediated RNA *trans*-splicing) technology, a patented spliceosome-mediated *trans*-splicing process owned by VIRxSYS, attempts to correct cellular damage caused by the formation of aberrant proteins by fixing or “reprogramming” defective pre-mRNA molecules so that only normal proteins are produced, even when a mutation is harbored and persists in the DNA. Naturally occurring transcription produces the 5' end of the pre-mRNA, while the investigator provides the downstream sequences closer to the 3' end of the molecule (Fig. 1.9).

Gene expression can be effectively knocked down by *trans*-splicing a stop codon into the endogenous RNA. Succinctly, a stop codon is a molecular stop sign that will result in the cessation of protein synthesis when the stop codon is encountered. Interestingly, the net result is the same as RNAi (RNA interference;

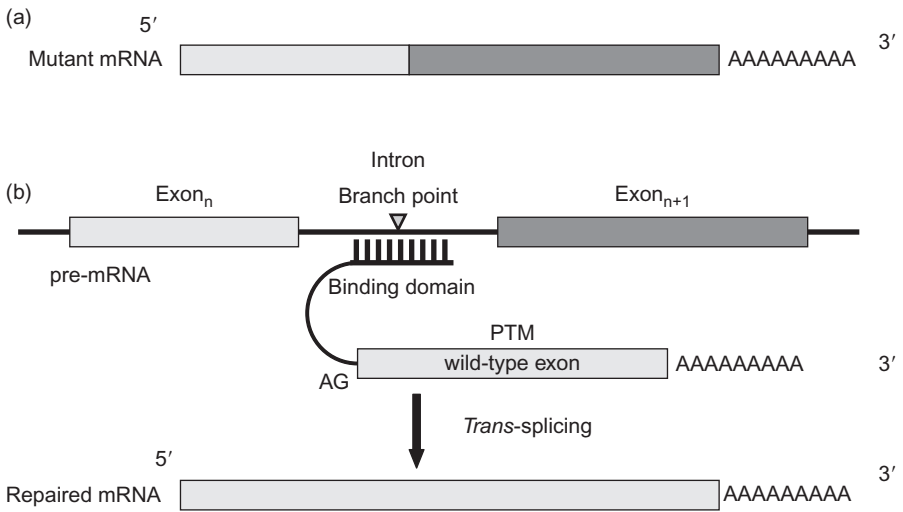


Figure 1.9 *Trans*-splicing joins together RNA from two different transcription units. (a). Mutant mRNA, capable of encoding a mutant protein. (b). *Trans*-splicing mRNA repair begins by targeting the pre-mRNA, which is produced in the cell from an endogenous gene locus. The RNA that will be spliced into the mutant pre-mRNA is referred to as the PTM (pre-*trans*-splicing molecule) and is likewise produced inside the cell following transfection of a PTM-encoding transcription vector. A splicing event involving the PTM binding domain and an intron that lies upstream of the mutation produces a repaired mRNA that is able to direct the synthesis of a wild-type protein. Notice that the PTM provides the requisite 'AG' splicing acceptor site. This is a superb example of an RNA therapy, rather than a gene therapy.

see Chapter 23); RNAi causes destruction of the mRNA template so that it cannot be translated, and *trans*-splicing a stop codon into an mRNA causes abrupt premature cessation of protein synthesis. Most truncated proteins in the cell have little or no function and will end up being recycled. *Trans*-splicing may well represent a low cost alternative to RNAi, particular when the exact RNAi sequences for gene silencing are elusive or the target cells are recalcitrant.

Overview of small RNAs

Small RNAs are the “new kid on the block” and the latest craze with respect to understanding the regulation of gene expression both *in vitro* and *in vivo*. With apparent ubiquity, these highly specialized molecules elicit profound changes that result in gene silencing, despite their remarkably small size of range of 19–26 nts, depending on their origin and function in the cell. Much of the early work in the realm of molecular biology missed these small but powerful molecules primarily because the original RNA isolation protocols, and later RNA isolation kits, were not designed to capture such small transcripts. It is now

evident that these small transcripts are of supreme regulatory importance, and kits for the isolation of small RNAs abound.

Small RNAs can be generically classified as one of several non-coding RNAs (ncRNA), meaning that the transcript does not encode a protein. Other examples of much larger ncRNAs include the previously discussed tRNA, rRNA, snoRNA, and Xist, which is a rather large sex-linked ncRNA that appears to be essential for the X-chromosome inactivation and the concomitant formation of Barr bodies observed in mammals (Brown *et al.*, 1992). Of particular interest in the realm of controlling gene expression are the two large families of small molecules classified as microRNA (miRNA) and short interfering RNA (siRNA). While the net result of the action of these two types of small RNA is the same (suppressing mRNA translatability), miRNAs are genome-encoded and tend to be conserved across species, whereas siRNAs result from precision cleavage of double-stranded RNA (dsRNA) from an exogenous source such as viral DNA or are laboratory-derived. Perhaps the best known application of small RNAs is that technology widely known as RNA interference (RNAi; see Chapter 23). It is important to note that siRNA-induced silencing is most often directed to the same locus, or closely related gene locus, from which the siRNA sequence was encoded while, in contrast, miRNAs tend to regulate completely different genes or families of genes. Other small RNAs are known to have a role in histone 3 methylation, resulting in gene silencing as a consequence of heterochromatin remodeling (Volpe *et al.*, 2002).

miRNA regulation of gene expression

All multicellular eukaryotes are believed to possess miRNAs and, as of this writing, more than 400 miRNAs are known in humans alone. Because of the undisputed importance of these regulatory molecules, several databases have been developed for cataloging miRNAs as they are discovered, including “The MicroRNA Registry” and “microRNA.org”, to name but a few.

miRNA “genes” are generally found in intergenic regions. Some miRNAs are organized into tandem repeat-like clusters, some are expressed independently, and yet other miRNAs are intron-associated. They tend to be highly concentrated in the vicinity of the centromere of each chromosome, an area once considered rather static and more much structural than functional. In a manner not unlike the synthesis of mRNA, transcription of miRNA DNA sequences produces large 5'-capped and 3'-polyadenylated precursor molecules known as pri-miRNAs that may contain one or several regions of intramolecular base-pairing referred to as hairpins or stem-loops. In turn pri-miRNAs are processed in the nucleus into pre-mRNAs and then further processed into double-stranded miRNAs. The ds miRNAs are then transported into the cytoplasm and dissociated into their constituent single strands by a helicase.



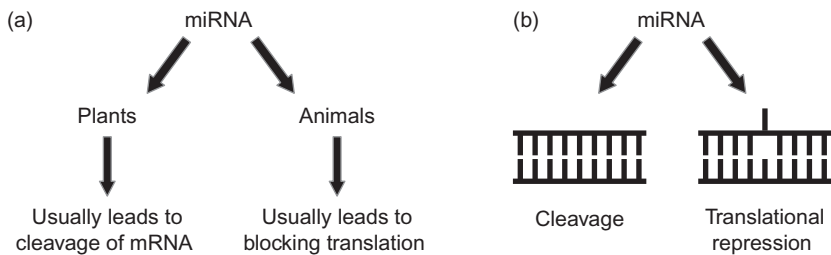


Figure 1.10 Regulatory roles for miRNA. Perfectly base-paired miRNAs direct the cleavage of the mRNA while mismatched miRNA molecules are able to suppress translation.

The antisense strand, i.e. the strand that is complementary to the mRNA, becomes associated with the RNA-induced silencing complex RISC, whereby direct interaction with the target mRNA becomes possible. Perfect base-pairing between mRNA and the RISC-associated miRNA results in transcript scission. Mismatching, in contrast, results in the repression of translation of mRNA (Fig. 1.10), the precise mechanism of which remains unclear. These same mechanisms of action apply to siRNA as well as miRNA, and it is widely believed that use of small RNAs in plants is part of a natural antiviral defense pathway. Small RNA activity may also embellish the innate immune system observed in higher animals.

Reference

- Ahmed, C.M.I., R.S. Chanda, N.D. Snow and B.S. Zain. (1982). The nucleotide sequence of mRNA for the M_r 19000 glycoprotein from early gene block III of adenovirus 2. *Gene* **20**, 339–346.
- Bag, J. (1991). mRNA and mRNP. In *Translation in Eukaryotes*. CRC Press, Ft. Lauderdale. pp. 71–79.
- Baltimore, D. (1970). Viral RNA-dependent DNA polymerase. *Nature* **226**, 1209–1211.
- Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol. Rev.* **44**, 175–205.
- Baumbach, L.L., F. Marashi, M. Plumb, G. Stein and J. Stein. (1984). Inhibition of DNA replication coordinately reduces cellular levels of H1 histone mRNAs requirement for protein synthesis. *Biochemistry* **23**, 1618–1625.
- Beelman, C.A. and R. Parker. (1995). Degradation of mRNA in eukaryotes. *Cell* **81**, 179–183.
- Benne, R., J. van den Berg, J.P. Brakenhoff, P. Sloof, J.H. Van Boom and M.C. Tromp. (1986). Major transcript of the coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819–826.
- Berget, S.M. and B.L. Robberson. (1986). U1, U2, and U4/U6 small nuclear ribonucleoproteins are required for *in vitro* splicing but not for polyadenylation. *Cell* **46**, 691–696.

- Birnsteil, M.L., M. Busslinger and K. Strub. (1985). Transcription termination and 3' end processing: the end is in site! *Cell* **41**, 349–359.
- Black, D.L. and J.A. Steitz. (1986). Pre-mRNA splicing *in vitro* requires intact U4/U6 small nuclear ribonucleoprotein. *Cell* **46**, 697–704.
- Blount, K.F. and O.C. Uhlenbeck. (2005). The structure–function dilemma of the hammerhead ribozyme. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 415–440.
- Brawerman, G. (1976). Characteristics and significance of the polyadenylate sequence in mammalian messenger RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **17**, 117–148.
- Breathnach, R. and P. Chambon. (1981). Organization and expression of eukaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**, 349–383.
- Brenner, S., F. Jacob and M. Meselson. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**, 576–581.
- Brown, C.J., B.D. Hendrich, J.L. Rupert, R.G. Lafrenière, Y. Xing, J. Lawrence and H. F. Willard. (1992). The human XIST gene: analysis of a 17kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**, 527–542.
- Buckingham, M.E., D. Caput, A. Cohen, R.G. Whalen and F. Gros. (1974). The synthesis and stability of cytoplasmic messenger RNA during myoblast differentiation in culture. *Proc. Natl. Acad. Sci.* **71**, 1466–1470.
- Cech, T.R., A.J. Zaugg and P.J. Grabowski. (1981). *In vitro* splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **27**, 487–496.
- Chen, C.Y.A. and A.B. Shyu. (1994). Selective degradation of early-response-gene mRNAs: Functional analysis of sequence features of the AU-rich elements. *Mol. Cell. Biol.* **14**, 8471–8482.
- Chen-Kiang, S., J.R. Nevins and J.E. Darnell. (1979). N-6-methyl-adenosine in adenovirus type 2 nuclear RNA is conserved in the formation of messenger RNA. *J. Mol. Biol.* **135**, 733–752.
- Coppola, J.A., A.S. Field and D.S. Luse. (1983). Promoter-proximal pausing by RNA polymerase II *in vitro*: transcripts shorter than 20 nucleotides are not capped. *Proc. Natl. Acad. Sci.* **80**, 1251–1255.
- Crick, F. (1957). On protein synthesis. Paper presented to the Society for Experimental Biology.
- Curtis, D., R. Lehmann and P.D. Zamore. (1995). Translational regulation in development. *Cell* **81**, 171–178.
- Davies, E., B.D. Fillingham, Y. Oto and S. Abe. (1991). Evidence for the existence of cytoskeleton-bound polysomes in plants. *Cell Biol. Intl. Reports* **15**, 973–981.
- Decker, C.J. and R. Parker. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: Evidence for a requirement for deadenylation. *Genes Dev.* **7**, 1632–1643.
- Decker, C.J. and R. Parker. (1994). Mechanisms of mRNA degradation in eukaryotes. *Trends Biochem. Sci.* **19**, 336–340.
- Dever, T.E. (2002). Gene-specific regulation by general translation factors. *Cell* **108**, 545–556.
- Dinesh-Kumar, S.P. and W.A. Miller. (1993). Control of start codon choice on a plant viral RNA encoding overlapping genes. *Plant Cell* **5**, 679–692.
- Doherty, E.A. and J.A. Doudna. (2000). Ribozyme structures and mechanisms. *Annu. Rev. Biochem.* **69**, 597–615.
- Dreyfuss, G., V.N. Kim and N. Kataoka. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* **3**, 195–205.

- Duyk, G.M., S. Kim, R.M. Myers and D.R. Cox. (1990). Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc. Natl. Acad. Sci.* **87**, 8995–8999.
- Edmonds, M., M.H. Vaughn, Jr. and H. Nakazato. (1971). Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly labeled poly-ribosomal RNA of HeLa cells: possible evidence for a precursor relationship. *Proc. Natl. Acad. Sci.* **68**, 1336–1340.
- Farrell, R.E. and C.L. Bassett. (2007). Multiple transcript initiation as a mechanism for regulating gene expression. In *Regulation of Gene Expression in Plants* (C.L. Bassett Ed.). Springer, New York, pp. 44–50.
- Fitzgerald, M. and T. Shenk. (1981). The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late sv40 mRNAs. *Cell* **24**, 251–260.
- Flavell, A.J., A. Cowie, S. Legon and R. Kamen. (1979). Multiple 5' terminal cap structures in late polyoma virus RNA. *Cell* **16**, 357–371.
- Florkiewicz, R.Z. and A. Sommer. (1989). Human basic fibroblast growth factor gene encodes four polypeptides: three initiate translation from non-AUG codons. *Proc. Natl. Acad. Sci.* **86**, 3978–3981.
- Fraser, N.W., J.R. Nevins, E. Ziff and J.E. Darnell. (1979). The major late adenovirus type-2 transcription unit: termination is downstream from the last poly(A) site. *J. Mol. Biol.* **129**, 643–656.
- Furuichi, Y. and A.J. Shatkin. (1989). Characterization of cap structures. *Methods Enzymol.* **180**, 164–177.
- Furuichi, Y., A. LaFiandra and A.J. Shatkin. (1977). 5'-terminal structure and mRNA stability. *Nature* **266**, 235–241.
- Futterer, J. and T. Hohn. (1996). Translation in plants – rules and exceptions. *Plant Mol. Biol.* **32**, 159–189.
- Gagliardi, D. and C.J. Leaver. (1999). Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J.* **18**, 3757–3766.
- Ghosh, P.K., V.B. Reddy, J. Swinscoe, P. Lebowitz and S.M. Weissman. (1978). Heterogeneity and 5' terminal structures of the late RNAs of simian virus 40. *J. Mol. Biol.* **126**, 813–846.
- Görlach, M. C.G. Burd and G. Dreyfus. (1994). The mRNA poly(A)-binding protein: Localization, abundance, and RNA-binding specificity. *Exp. Cell Res.* **211**, 400–407.
- Gros, F., H. Hiatt, W. Gilbert, C.G. Kurland, R.W. Risebrough and J.D. Watson. (1961). Unstable ribonucleic acid revealed by pulse labelling of *Escherichia coli*. *Nature* **190**, 581–585.
- Guerrier-Takada, C. and S. Altman. (1984). Catalytic activity of an RNA molecule prepared by transcription *in vitro*. *Science* **223**, 285–286.
- Hall, B.D. and S. Spiegelman. (1961). Sequence complementarity of T2 DNA and T2-specific RNA. *Proc. Natl. Acad. Sci.* **47**, 137–146.
- Hamm, J. and I.W. Mattaj. (1990). Structural analysis of U small nuclear ribonucleoproteins by *in vitro* assembly. *Methods Enzymol.* **181**, 273–284.
- Hann, S.R., M.W. King, D.L. Bentley, W.C. Anderson and R.N. Eisenman. (1988). A non-AUG translational initiation in *c-myc* exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. *Cell* **52**, 185–195.
- Hashimoto, C. and J.A. Steitz. (1986). A small nuclear ribonucleoprotein associates with AAUAAA polyadenylation signal *in vitro*. *Cell* **45**, 581–591.

- Hastings, K. (2005). SL trans-splicing: easy come or easy go? *Trends Genet.* **21**, 240–247.
- Hayes, R., J. Kudla and W. Gruissem. (1999). Degradation of chloroplast mRNA: the role of polyadenylation. *Trends Biochem. Sci.* **24**, 199–202.
- Herr, A.J., M.B. Jensen, T. Dalmay and D.C. Baulcombe. (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120.
- Hewlett, M.J., J.K. Rose and D. Baltimore. (1976). 5'-terminal structure of poliovirus polyribosomal RNA is pUp.. *Proc. Natl. Acad. Sci.* **73**, 327–330.
- Hirose, Y. and J.L. Manley. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev.* **14**, 1415–1429.
- Huez, G., G. Marbaix, E. Hubert, M. Leclercq, U. Nudel, H. Soreq, R. Salomon, R. Lebleu, M. Revel and U.Z. Littauer. (1974). Role of the polyadenylate segment in the translation of globin messenger RNA in *Xenopus* oocytes. *Proc. Natl. Acad. Sci.* **71**, 3143–3146.
- Hunt, A.G. (2007). Messenger RNA 3'-end formation and the regulation of gene expression. In *Regulation of Gene Expression in Plants* (C.L. Bassett Ed.). Springer, New York. pp. 101–122.
- Jackson, R.J. and N. Standart. (1990). Do the poly(A) tail and 3' untranslated region control mRNA translation? *Cell* **62**, 15–24.
- Jacob, F. and J. Monod. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Molec. Biol.* **3**, 318–356.
- Jung, A., A.E. Sippel, M. Grez and G. Schutz. (1980). Exons encode functional and structural units of chicken lysozyme. *Proc. Natl. Acad. Sci.* **77**, 5759–5763.
- Kanno, T., B. Huettel, M.F. Mette, W. Aufsatz, E. Jalignot, L. Daxinger, D.P. Kriel, M. Matzke and A.J.M. Matzke. (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nature Genet.* **37**, 761–765.
- Kass, S., K. Tyc, J.A. Steitz and B. Sollner-Webb. (1990). The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* **60**, 897–908.
- Kim, M., W. Canio, S. Kessler and N. Sinha. (2001). Developmental changes due to long-distance movement of a homeobox fusion transcript in tomato. *Science* **29**, 287–289.
- Knight, E., E.D. Anton, D. Fahey, B.K. Friedland and G.J. Jonak. (1985). Interferon regulates *c-myc* expression in Daudi cells at the post-transcriptional level. *Proc. Natl. Acad. Sci.* **82**, 1151–1154.
- Kozak, M. (1978). How do eukaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15**, 1109–1123.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation in eukaryotes. *Cell* **44**, 283–292.
- Kozak, M. (1989). The scanning model for translation: an update. *J. Cell Biol.* **108**, 229–241.
- Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci.* **87**, 8301–8305.
- Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867–19870.
- Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187–208.
- Kravchenko, J.E., I.B. Rogozin, E.V. Koonin and P.M. Chumakov. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondria origin. *Nature* **436**, 735–739.

- Kruger, K., P.J. Grabowski, A.J. Zaug, J. Sands, D.E. Gottschling and T.R. Cech. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31**, 147–157.
- Kubo, T., T. Wada, Y. Yamaguchi, A. Shimizu and H. Handa. (2006). Knock-down of 25 kDa subunit of cleavage factor IM in HeLa cells alters alternative polyadenylation within 3'-UTRs. *Nucleic Acids Res.* **34**, 6264–6271.
- Kuge, H. and J.D. Richter. (1995). Cytoplasmic 3' poly(A) addition induces 5' cap ribose methylation: implications for translational control of maternal mRNA. *EMBO J.* **14**, 6301–6310.
- Kurreck, J. (2003). Antisense technologies. Improvement through novel chemical modifications. *European J. Biochem.* **270**, 1628–1644.
- Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nature Genet.* **30**, 363–364.
- Le Hir, H., M.J. Moore and L.E. Maquat. (2000). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. *Genes Dev.* **14**, 1098–1108.
- Lykke-Andersen, J., M.D. Shu and J.A. Steitz. (2000). Human Upf proteins target mRNP for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103**, 1121–1131.
- Leibold, E.A. and H.N. Munro. (1988). Cytoplasmic proteins binds *in vitro* to a highly conserved sequence in the 5' untranslated region of ferritin heavy and light-subunit mRNAs. *Proc. Natl. Acad. Sci.* **85**, 2171–2175.
- Lenk, R., L. Ransom, Y. Kaufman and S. Penman. (1977). A cytoskeletal structure with associated polyribosomes obtained from HeLa cells. *Cell* **10**, 67–78.
- Lewin, B. (Ed.). (2008). *Genes IX*. Jones and Bartlett Publishers, Sudbury, MA.
- Lim, L. and E.S. Cannelakis. (1970). Adenine-rich polymer associated with rabbit reticulocyte messenger RNA. *Nature* **227**, 710–712.
- Lisitsky, I., P. Klaff and G. Schuster. (1997). Blocking polyadenylation of mRNA in the chloroplast inhibits its degradation. *Plant J.* **12**, 1173–1178.
- Louzada, E.S. (2007). Alternative processing as a Mechanism for regulating gene expression. In *Regulation of Gene Expression in Plants* (C.L. Bassett Ed.). Springer, New York. pp. 67–100.
- Mandal, M. and R.R. Breaker. (2004). Gene regulation by riboswitches. *Nature Rev. Molec. Cell Biol.* **5**, 451–463.
- Manley, J.L. and N.J. Proudfoot. (1994). RNA 3' ends – formation and function. *Genes Dev.* **8**, 259–264.
- Manley, J.L., H. Yu and L. Ryner. (1985). RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation *in vitro*. *Mol. Cell. Biol.* **5**, 373–379.
- Maquat, L.E. (1997). RNA export from the nucleus. In *mRNA Metabolism and Post-Transcriptional Gene Regulation* (J.B. Harford and D.R. Morris, Eds). Wiley-Liss, New York, NY. pp. 107–126.
- Marbaix, G., G. Huez, A. Burny, Y. Cleuter, E. Hubert, M. Leclercq, H. Chantrenne, H.Soreq, U.Nudel and Z.Littauer. (1975). Absence of polyadenylate segment in globin messenger RNA accelerates its degradation in *Xenopus* oocytes. *Proc. Natl. Acad. Sci.* **72**, 3065.
- Marzluff, W.F. and R.C.C. Huang. (1984). Transcription of RNA in isolated nuclei. In *Transcription and Translation: A Practical Approach* (B.D. Hames and S.J. Higgins, Eds). IRL Press, Washington, DC. pp. 89–129.

- Montange, R.K. and R.T. Batey. (2008). Riboswitches: emerging themes in RNA structure and function. *Annual Rev. Biophys.* **37**, 117–133.
- Mount, S.M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**, 459–472.
- Nagaswamy, U., M. Larios-Sanz, J. Hury, S. Collins, Z. Zhang, Q. Zhao and G.E. Fox. (2002). NCIR: A database of non-canonical interactions in known RNA structures. *Nucleic Acids Res.* **29**, 395–397.
- Nahvi, A., N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown and R.R. Breaker. (2002). Genetic control by metabolite binding mRNA. *Chem. Biol.* **9**, 1043–1049.
- Nevins, J.R. and J.E. Darnell. (1978). Steps in the processing of Ad2 mRNA: poly(A)⁺ nuclear sequences are conserved and poly (A) addition precedes splicing. *Cell* **15**, 1477–1493.
- Nomoto, A., Y.F. Lee and E. Wimmer. (1976). The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np.. *Proc. Natl. Acad. Sci.* **73**, 375–380.
- O'Hara, E.B., J.A. Chekanova, C.A. Ingle, Z.R. Kushner, E. Peters and S.R. Kushner. (1995). Polyadenylation helps regulate mRNA decay in *E. coli*. *Proc. Natl. Acad. Sci.* **92**, 1807–1811.
- Onodera, Y., J.R. Haag, T. Ream, P.C. Nunes, O. Pontes and C.S. Pikaard. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622.
- Pain, V. (1996). Initiation of protein synthesis in eukaryotic cells. *Eur. J. Biochem.* **236**, 747–771.
- Piñol-Roma, S. and G. Dreyfus. (1993). hnRNP proteins: Localization and transport between the nucleus and the cytoplasm. *Trends Cell Biol.* **7**, 151–155.
- Preobrazhensky, A.A. and A.S. Spirin. (1978). Informosomes and their protein components: the present state of knowledge. *Prog. Nucleic Acid Res. Mol. Biol.* **21**, 1–38.
- Proudfoot, N.J. and G.G. Brownlee. (1976). 3' Noncoding region sequences in eucaryotic messenger RNA. *Nature* **263**, 211–214.
- Reddy, R., T.S. Ro-Choi, D. Henning and H. Busch. (1974). Primary sequence of U-1 nuclear ribonucleic acid of Novikoff hepatoma ascites cells. *J. Biol. Chem.* **249**, 6486–6494.
- Richter, D. (1983). Vasopressin and oxytocin are expressed as polyproteins. *Trends Biochem. Sci.* **8**, 278–291.
- Roeder, R.G. (1976). In “RNA Polymerase” (Losick R. and M. Chamberlain Eds). (1976). *RNA Polymerase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. pp. 285–329.
- Sachs, A.B. and R.D. Kornberg. (1990). Purification and characterization of polyadenylate-binding protein. *Methods Enzymol.* **181**, 332–352.
- Schuster, G., I. Lisitsky and P. Klaff. (1999). Update on chloroplast molecular biology: polyadenylation and degradation of mRNA in the chloroplast. *Plant Physiol.* **120**, 937–944.
- Shatkin, A.J. (1976). Capping of eucaryotic mRNAs. *Cell* **9**, 645–653.
- Shaw, G. and R. Kamen. (1986). A conserved AU sequence from 3' untranslated region of GM-CSF mRNA mediates selective messenger RNA degradation. *Cell* **46**, 659–667.
- Shyu, A.B., J.G. Belasco and M.E. Greenberg. (1991). Two distinct destabilizing elements in the *c-fos* message trigger deadenylation as a first step in rapid mRNA decay. *Genes Dev.* **5**, 221–231.
- Spirin, A.S. (1994). Storage of messenger RNA in eukaryotes: Envelopment with protein, translational barrier at 5' side, or conformational masking by 3' side. *Mol. Reprod. Dev.* **38**, 107–117.

- Standart, N. and R.J. Jackson. (1994). Regulation of translation by specific protein/mRNA interactions. *Biochimie* 76, 867–879.
- Taub, R., I. Kirsch, C. Morton, G. Lenoir, D. Swan, S. Tronick, S. Aaronson and P. Leder. (1982). Translocation of the *c-myc* oncogene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc. Natl. Acad. Sci.* 79, 7837–7841.
- Tavernier, J., R. Derynck and W. Fiers. (1981). Evidence for a unique human fibroblast interferon (IFN- β_1) chromosomal gene, devoid of intervening sequences. *Nucleic Acids Res.* 9(3), 461–471.
- Temin, H. and S. Mizutani. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211–1213.
- Tosi, M., R.A. Young, O. Hagenbuchle and U. Schibler. (1981). Multiple polyadenylation sites in a mouse α -amylase gene. *Nucleic Acids Res.* 9, 2313–2323.
- Triesman, R., N.J. Proudfoot, M. Shander and T. Maniatis. (1982). A single-base change at a splice site in a β^0 -thalassemic gene causes abnormal RNA splicing. *Cell* 29, 903–911.
- Volkin, E. and L. Astrachan. (1956). Phosphorus incorporation of *E. coli* ribonucleic acid after infection with bacteriophage T2. *Virology* 2, 146–161.
- Volpe, T.A., C. Kidner, I.M. Hall, G. Teng, S.I. Grewel and R.A. Martienssen. (2002). Regulation of heterchromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833–1837.
- Wahle, E. and U. Ruesegger. (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* 23, 277–295.
- Watson, J.D., T.A. Baker, S.P. Bell, A. Gann, M. Levine and R. Losick. (2008). *Molecular Biology of the Gene*, 6th edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY pp. 472.
- Will, C.L. and B.J. Dolnick. (1989). 5-Fluorouracil inhibits dihydrofolate reductase precursor mRNA processing and/or nuclear mRNA stability in methotrexate-resistant KB cells. *J. Biol. Chem.* 264, 21413–21421.
- Wilson, T. and R. Treisman. (1988). Removal of poly(A) and consequent degradation of *c-fos* mRNA facilitated by 3' AU-rich sequences. *Nature* 336, 396–399.
- Winkler, W., A. Nahvi and R.R. Breaker. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952–956.
- Winkler, W.C., S. Cohen-Chalamish and R.R. Breaker. (2002). An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci.* 99(25), 15908–15913.
- Woolin, S.L. and P. Walter. (1988). Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J.* 7, 3559–3569.
- Zamenhof, S., G. Brawermann and E. Chargaff. (1952). On the desoxypentose nucleic acids from several organisms. *Biochim. Biophys. Acta* 9, 402–405.
- Zeevi, M., J.R. Nevins and J.E. Darnell, Jr. (1982). Newly formed mRNA lacking polyadenylic acid enters the cytoplasm and the polyribosomes but has a shorter half-life in the absence of polyadenylic acid. *Mol. Cell Biol.* 2, 517–525.
- Zhang, X., I.R. Henderson, C. Luc, P.J. Green and S.E. Jacobsen. (2007). Role of RNA polymerase IV in plant small RNA metabolism. *Proc. Natl. Acad. Sci.* 104, 4536–4541.

2 RNA Isolation Strategies

Rationale 45	Simultaneous isolation of RNA and DNA 68
Goals in the purification of RNA 47	Protocol: simultaneous isolation of RNA and DNA 69
Lysis buffer formulations 49	The word on kits 71
Gentle lysis buffers 49	Silica technology 72
Protocol: isolation of cytoplasmic RNA by gentle hypotonic lysis 51	Isolation of cytoplasmic RNA on a silica column 73
Chaotropic lysis buffers 56	Affinity matrices 73
Isolation of RNA with guanidinium buffers 57	Other methods 74
Guanidinium–acid–phenol extraction techniques 58	Protocol: rapid isolation of RNA with SDS and potassium acetate reagents 74
Protocol: guanidinium–acid–phenol extraction 59	Protocol: isolation of prokaryotic RNA 75
Density gradient centrifugation 60	Protocol: isolation of RNA from yeast 76
Cesium chloride 61	Short- and long-term storage of purified RNA 78
Protocol: cesium chloride (CsCl) gradients 61	References 79
Cesium trifluoroacetate (CsTFA) 65	
Protocol: cesium trifluoroacetate (CsTFA) gradients 65	

Rationale

RNA is chemically and biologically more labile than DNA, particularly at elevated temperatures ($>65^{\circ}$) and in the presence of alkali. These intrinsic handling difficulties are further compounded by aggressive activities of a variety of resilient ribonucleases (RNase), the apparent ubiquity of which is undisputed. The isolation of RNA from biological sources enriched in RNase is especially challenging. These factors mandate rapid recovery of RNA with reagents that are free of RNase contamination. Likewise, extraction tactics must be formulated to include sufficient management of endogenous RNase activity upon disruption of the organelles and vacuoles that normally sequester these enzymes. Failure to scrupulously observe procedures for eliminating potential sources of RNase contamination will almost always yield a useless sample of degraded RNA. Further, even trace quantities of contaminants carried over from some isolation procedures may compromise the efficiency of subsequent reactions.

Due to their naturally short half-life, experiments involving RNA are most judiciously planned around the availability of cell cultures or tissue samples and the actual date of RNA isolation.

Efficient methodologies have been empirically derived to accommodate the expedient isolation of RNA, techniques that should be scrutinized and refined continuously. In general, these methods yield cytoplasmic RNA, nuclear RNA, or mixtures of both, commonly known as total or cellular RNA. Protocols for the isolation of RNA begin with cellular lysis mediated by buffers that typically fall into one of two categories: (1) those consisting of harsh chaotropic agents such as one of the guanidinium salts¹, sodium dodecyl sulfate (SDS), N-laurylsarcosine (sarcosyl), urea, phenol, or chloroform, which disrupt the plasma membrane and subcellular organelles and which simultaneously inactivate RNase; and (2) those that gently solubilize the plasma membrane while maintaining nuclear and other organelle integrity, such as the non-ionic, hypotonic lysis buffers²; organelles and cellular debris are then removed from the lysate by differential centrifugation. The best known and most efficient lysis buffers in this category contain Nonidet™ P-40 (NP-40). This detergent, however, is no longer commercially available. IGEPAL® CA-630 (Sigma Cat. I-3021), which is chemically indistinguishable from NP-40, can be substituted. It is worth noting that a seemingly endless list of permutations on a few fundamental RNA extraction techniques exist; for example, some techniques support isolation of poly(A)⁺ material directly from a cellular lysate without prior purification of total RNA. As far as which RNA extraction procedure to use, it is incumbent upon the investigator to always think two steps ahead and ask “What is to be done with this RNA after it has been purified?”

RNA isolated from the nucleus or cytoplasm by direct cell lysis is known as *steady-state* RNA; it represents the final accumulation of the RNA in the cell or a subcellular compartment. Whereas the abundance of specific steady-state RNA species may certainly be interpreted as an indication of how specific genes are modulated, it furnishes information neither about the *rate* at which these RNA molecules are transcribed nor about their stability or half-life in the cell. This is the major disadvantage of evaluating the steady-state RNA alone. The cellular biochemistry responds to environmental change not only by altering the rate of transcription, but also through the processing efficiency of precursor mRNA, the efficiency of nucleocytoplasmic transport, the stability of RNA in the cytoplasm, and the translatability of salient messages. Protocols for the analysis of steady-state transcripts are presented in this chapter, while protocols for the analysis of nuclear RNA, including transcription rate assays, are presented in Chapter 16.

¹Guanidinium thiocyanate = GTC; guanidinium hydrochloride = G-HCl.

²While osmotic lysis is one of the most gentle methods for cell disruption, it does nothing when carbohydrate-rich cell walls are present, as with certain bacteria, fungi, and plant cells. Appropriate steps must be taken to break through the cell wall and access the cellular contents.

Goals in the purification of RNA

Any worthwhile strategy for RNA purification must address and achieve specific goals if data derived from the final RNA preparation are to be meaningful. Before describing the goals associated with RNA isolation methods it is worth defining the subpopulations of RNA that are commonly of experimental interest (Table 2.1). It is incumbent upon the investigator to consider which population of RNA will result from a particular isolation procedure before picking up the micropipettors.

Goal 1: Select an appropriate method of membrane solubilization

The first decision to factor into the design of an RNA isolation strategy is the method of cellular disruption, and it is based on which population of RNA or subcellular compartment the investigator wishes to study. For example, an investigator may wish to determine whether the observed modulation of gene expression in a model system is regulated transcriptionally or by some post-transcriptional event(s). In such an instance, selection of the method of cellular disruption and subsequent RNA isolation must permit analysis of the nuclear abundance of salient transcripts independently of the cytoplasmic abundance of the same species.

The method of cell lysis will determine the extent of subcellular disruption of the sample. For example, a lysis buffer that is used successfully with tissue culture cells may be entirely inappropriate for whole tissue samples. Just as important, the method by which membrane solubilization is accomplished will dictate whether additional steps will be required to remove DNA from the RNA preparation, and whether compartmentalized nuclear RNA and cytoplasmic RNA species can be purified independently of one another. While DNA can be purged from an RNA preparation with minimal fanfare, it is not possible to determine the relative contribution of RNA from the nucleus and the cytoplasm

Table 2.1 Subpopulations of Eukaryotic RNA

mRNA Subpopulation	Composition
Total cellular RNA	All RNA from the cell
Total cytoplasmic RNA	Cytoplasmic RNA, excluding mtRNA
Nuclear RNA	Transcripts (mature and immature) isolated directly from intact nuclei.
mRNA	Transcripts that have matured and been exported to the cytoplasm. This subpopulation includes both poly(A) ⁺ and poly(A) ⁻ transcripts.
Cellular Poly(A) ⁺ RNA	Nuclear and cytoplasmic polyadenylated transcripts
Cytoplasmic Poly(A) ⁺ RNA	Cytoplasmic polyadenylated mRNA
Poly(A) ⁻ RNA	Non-polyadenylated RNA (mostly rRNA and tRNA)
mtRNA	Mitochondrial RNA (all types)
cpRNA	Chloroplast RNA (all types; plants only)

once RNA from these two subcellular compartments have co-purified. A particular lysis procedure must likewise demonstrate compatibility with ensuing protocols once the RNA has been recovered. Always think two steps ahead; the proper method of solubilization is dependent on the plans for the RNA after purification and the questions being asked in a particular study.

Goal 2: Ensure total inhibition of nuclease activity

The imperative for controlling nuclease activity should be abundantly clear, if not from personal experience, then from the discussion in Chapter 7. This includes purging RNase from reagents, glassware, consumable plasticware, and equipment and controlling such activity in a cell lysate. Although some lysis reagents inhibit nuclease activity in their own right, other lysis reagents require additional nuclease inhibitors to safeguard the RNA during the isolation procedure. Steps for the inhibition or elimination of RNase activity must, first and foremost, demonstrate compatibility with the lysis buffer.

Goal 3: Select a method for deproteinization of the sample

The complete removal of protein from a cellular lysate is of paramount importance in the isolation of both DNA and RNA. Meticulous attention to this detail is required for accurate quantification and precision in hybridization. For example, the restriction of double-stranded DNA (dsDNA) and the action of DNA- and RNA ligase can be inhibited by carry-over protein, especially histone proteins. Further, reverse transcription (Chapter 17) and amplification by PCR (Chapter 18) are inhibited strongly by the use of “dirty” RNA. Removal of proteins may be accomplished by:

- a. Digestion of the sample with the enzyme proteinase K.
- b. Repeated extraction with mixtures of organic solvents such as phenol and chloroform.
- c. Solubilization in guanidinium buffers.
- d. Physical separation by (silica) column chromatography.
- e. Salting-out of proteins.
- f. Any combination of the above.

Any procedure for deproteinization is itself a means of controlling RNase activity. It is important to recognize that the RNA in the sample will be susceptible to nuclease degradation once the protein denaturant or inhibitor is removed; RNA is never more susceptible to nuclease attack than when it is dissolved in simple, non-denaturing aqueous buffer such as TE buffer or nuclease-free water.

Goal 4: Select a method for nucleic acid concentration

This is the final step in most RNA purification schemes. The most versatile method for concentrating nucleic acids is precipitation using various combinations of salt and alcohol (Table 2.2). For example, one common method is to add 0.1 volume of 3M sodium acetate (pH 5.2) to a nucleic acid sample, followed by the addition of 2.5 volumes of 95–100% ethanol. Nucleic acids

and the salt that drives their precipitation form complexes that have greatly reduced solubility in ethanol and isopropanol. Further, the rate of precipitation using various salt and alcohol combinations is temperature-dependent. Unlike the dramatic precipitation of genomic DNA, the precipitation of RNA is much more dignified, often requiring longer incubation periods at -20° to ensure complete recovery, especially when using the non-ionic lysis procedure described below. This phenomenon is a direct function of genome size and complexity that an organism exhibits, compared to the relatively low complexity of cellular RNA. Other procedures for concentrating nucleic acid samples³, including the use of commercially available concentrating devices, dialysis, and centrifugation under vacuum, are not as frequently employed because the degree of success is much more dependent on the skill of the investigator than with simple salt and alcohol concentration. Finally, the use of the newer silica-based technologies often precludes the need to add any salt or alcohol because the RNA can be eluted from the column in concentrated form and ready for immediate use.

Goal 5: Select the proper storage conditions for purified samples of RNA

Because of the naturally labile character of RNA, improper storage of excellent RNA samples will often result in degradation in a relatively short time. There are many opinions as to the proper temperature, buffer, and storage form for both RNA and DNA. A discussion of storage considerations and options appears at the end of this chapter.

Lysis buffer formulations

Everyone has a favorite method for the isolation of the RNA, and there are numerous permutations on the protocols that accompany many of the commercially available products for RNA isolation. In general, however, all of these methodologies fall into one of two major categories, depending on the degree of cellular disruption and concomitant RNase inhibition. These categories are referred to herein as the gentle lysis buffers and the harsh lysis buffers. Each has a role in the investigation of gene expression, and each has limitations, as described in the following sections. The most important thing to remember is that there is no one right way to extract RNA from cells and tissues and the only wrong way of doing things is to permit the introduction of RNase activity.

Gentle lysis buffers

Cellular lysis mediated by non-ionic, hypotonic buffer is not disruptive to most subcellular organelles. The inclusion of non-ionic detergents such as NP-40

³An older procedure for concentration of nucleic acids, repeated extraction with butanol, is not recommended because of the inherent difficulties associated with the complete removal of butanol.

Table 2.2 Salt (A) and Alcohol (B) Combinations for Nucleic Acid* Precipitation

A. Salt	Typical stock concentration	Amount required	Final concentration
NaOAc	3 M, pH 5.2	0.1 volume	300 mM
NaCl	1 M	0.1 volume	200 mM
LiCl**	8 M	0.1 volume	800 mM
NH ₄ OAc***	10 M	0.2 volume	2 M
KOAc	2.5 M	0.1 volume	250 mM
Glycogen****	10 mg/ml	0.01 volume	100 µg/ml

B. Alcohol	Volume required after addition of salt
Ethanol (95–100%)	2.2–2.5 volumes
Isopropanol (100%)	0.6–1.0 volume

*Precipitation thresholds: DNA 50 ng/ml, RNA 100 ng/ml. If the concentration falls below the threshold, precipitation will be very inefficient unless a carrier is added.

**Lithium chloride does not coprecipitate with nucleic acids. As with NaCl, it is very soluble in ethanol. Avoid LiCl if RNA is to be reverse transcribed unless the RNA will be double-precipitated in order to remove the excess LiCl.

***Preferred salt for precipitating cDNA after synthesis is ammonium acetate because it does not precipitate unincorporated dNTPs.

****The judicious use of glycogen is recommended because excessive quantities will inhibit a number of *in vitro* reactions.

(Favaloro *et al.*, 1980) and MgCl₂ in the lysis buffer facilitates plasma membrane solubilization while maintaining nuclear integrity. Thus, initial lysis of cells with NP-40-containing buffer precludes the mixing of nuclear and cytoplasmic RNA from the onset. In such protocols, intact nuclei, large organelles, and cellular debris are easily removed by differential centrifugation. The resulting cytosolic supernatant is rich in cytoplasmic RNA and proteins, the latter being easily removed by a series of extractions with mixtures of organic solvents such as phenol and chloroform or, preferably, by silica column chromatography. The obvious advantage of this isolation strategy is that the RNA precipitated at the end of the procedure represents only the cytoplasmic population. This may be especially meaningful in the exploration of the level of gene regulation.

A disadvantage of using this approach is that the lysis buffer alone is not sufficiently chaotropic to fully inhibit RNase activity. Keep in mind that upon cellular lysis, normally sequestered RNases are suddenly liberated, and their activity will greatly compromise the integrity of the RNA even as the investigator is working diligently to purify it. It may be helpful, but is necessary, to add some type of RNase inhibitor to the lysis buffer *just prior to use* in order to control nuclease activity. In this laboratory very high-quality RNA is recovered when the sample is maintained on ice at all times unless the protocol specifically dictates otherwise. It is also very helpful to use ice-cold reagents and tubes that have been pre-chilled. Should one wish to take advantage of the disruptive

nature of the guanidinium-based isolation procedures discussed below and still maintain the subcellular compartmentalization of RNA, one worthwhile strategy is to begin the isolation procedure with a non-ionic lysis buffer, recover intact nuclei, and then lyse the nuclei with guanidinium buffer; one may proceed to purify the nuclear (or cytoplasmic) RNA as if working with intact cells. This approach is particularly suited for the isolation of nuclear RNA. Protocols describing the use of guanidinium buffers are described here; for a more complete discussion of the isolation of nuclear RNA, see Chapter 16.

Protocol: isolation of cytoplasmic RNA by gentle hypotonic lysis

This procedure is based on non-ionic, hypotonic solubilization of the plasma membrane while maintaining nuclear integrity⁴. It may be used for the isolation of cytoplasmic RNA as described here and is likewise used for the isolation of intact nuclei and accompanying purification of nuclear RNA, as described in Chapter 16. On a personal note, this procedure is utilized routinely using ice-cold reagents and without any RNase inhibitors, and renders very high quality RNA. An in-house modification of this valuable technique is presented here (Fig. 2.1). If RNase inhibitors are to be used they must remain with the lysate until the phenol:chloroform mixtures are introduced into the system to remove the protein. Finally, the RNA is recovered by salt and alcohol precipitation of the RNA, a widely used preparative technique. Whereas this method is faster than the older guanidinium-based methods requiring ultracentrifugation, it may appear to be more time-consuming than the newer acid-phenol methods (below). This is due, in part, to the requirement for gentle handling of the sample to preclude nuclear breakage, which, to a limited extent, *may* put the RNA molecules at slightly greater risk of RNase degradation. RNA prepared according to this method is quite adequate for most standard applications. This procedure works best when starting with cells grown in tissue culture, as the lysis buffer does not disrupt whole tissue samples as efficiently as cells harvested from tissue culture. The protocol may be adapted for whole tissue samples when combined with gentle Dounce homogenization to prevent serious damage to organellar membranes.

In advance: Preparation of extraction buffer

An adequate volume of extraction buffer should be freshly prepared before each use, as the oxidation of phenol can greatly compromise the quality of an RNA preparation (Fig. 2.2).

1. Melt redistilled (molecular biology grade) phenol at 60° and, as shown in Fig. 2.2A, saturate it by mixing the phenol with an equal volume of Tris-SDS buffer:
100 mM NaCl
1 mM EDTA

⁴The investigator may wish to optimize empirically the lysis conditions because cells and organelles exhibit different sensitivities to detergents, salts, and RNase inhibitors.

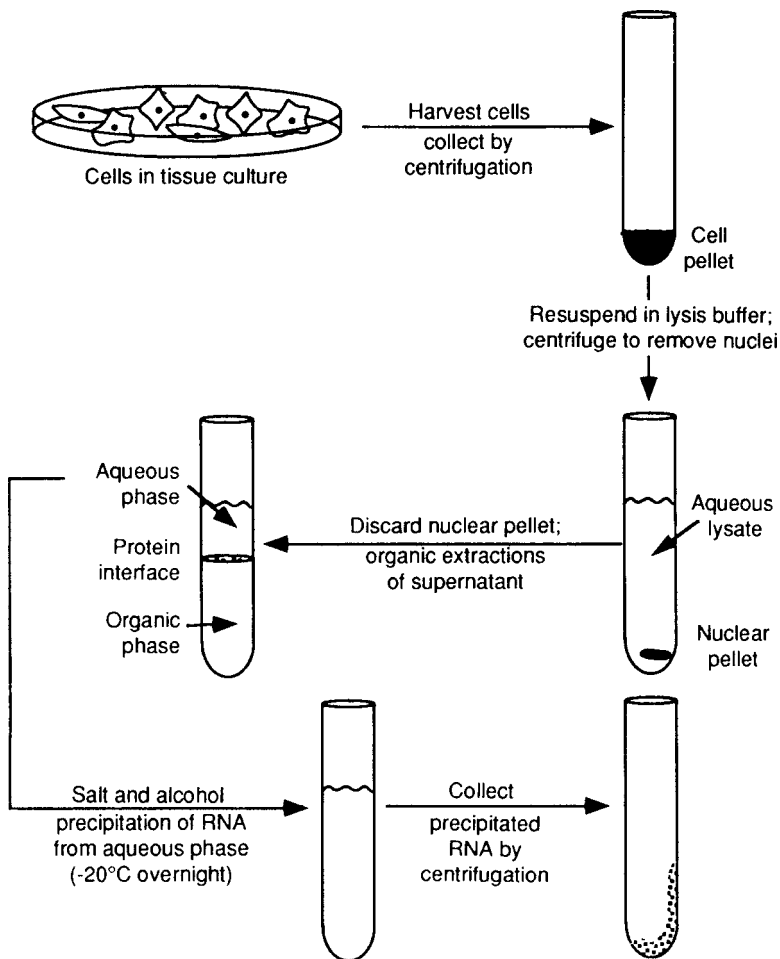


Figure 2.1 Preparation of total cytoplasmic RNA by gentle hypotonic. This method is best suited for cells grown in culture; soft tissue samples may require a gentle, non-aggressive homogenization step.

10mM Tris, pH 8.5
0.5% SDS⁵

- To the resulting Tris-SDS/phenol mixture add an equal volume of a mixture of chloroform:isoamyl alcohol (24:1) (Fig. 2.2B). The resulting reagent is known as the extraction buffer (Fig. 2.2C).
- Add 8-hydroxyquinoline to the extraction buffer (Fig. 2.2D) to a final concentration of 0.1% (w/v). Allow the phases to separate at 4°.

⁵The inclusion of SDS ensures the quantitative recovery of poly(A)⁺ RNA in the aqueous phase during extraction with organic buffer. Saturation of phenol with Tris-SDS buffer is desirable also because this partitioning phenomenon in the presence of SDS is pH-independent. Be sure to autoclave this buffer, or at least the stock solutions, *before* adding SDS to complete this reagent.

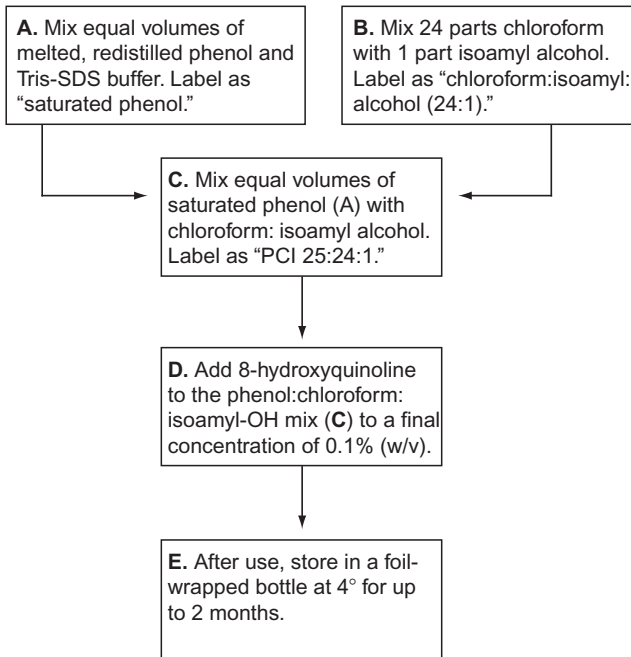


Figure 2.2 Method of preparation of the organic extraction mixture used for protein removal. Saturation of phenol with Tris/SDS buffer, rather than Tris alone, and the inclusion of 8-hydroxyquinoline renders a superior formulation for the deproteinization of the cytosol. This buffer can be substituted in most applications when the use of phenol:chloroform:isoamyl alcohol is dictated, and can be used in both DNA and RNA isolation procedures.

For example: Mix 20 ml of melted, redistilled phenol with 20 ml (an equal volume) of Tris/SDS buffer. To the resulting 40 ml of Tris/SDS-saturated phenol, add 40 ml (an equal volume) of chloroform:isoamyl alcohol, previously mixed in a ratio of 24:1. This is the extracting buffer, the stability of which is enhanced by the addition of 0.08 g of 8-hydroxyquinoline (final concentration of 0.1% w/v). This buffer is conveniently prepared and stored in a 100 ml glass bottle and has a shelf life of about 4–8 weeks at 4°. Ideally, the bottle should be wrapped in foil when not in use, especially if the refrigerator is opened and closed continuously throughout the day.

CAUTION: Be sure to handle all organic solvents with proper skin and eye protection, in a chemical fume hood, and be sure to dispose of organic waste according to departmental regulations.

RNA isolation

1. Wear gloves!
2. Harvest cells from tissue culture and collect by centrifugation in a suitable tube, pre-chilled on ice. Note: For trypsinization protocol, see Appendix J.

3. Gently, and completely, resuspend the cell pellet in ice cold lysis buffer:
 - 140 mM NaCl
 - 1.5 mM MgCl₂
 - 10 mM Tris-Cl, pH 8.5
 - 0.5% NP-40⁶

Use at least 100 μ l lysis buffer/10⁶ cells.

4. Incubate on ice for 5 min. Periodic *gentle* inversion of the tube may facilitate cell lysis. Incubate on ice for 5 min.

Note 1: The extent of cell lysis may be assessed by examining a 1- μ l aliquot of the lysate on a microscope slide for the presence of intact cells versus free nuclei.

Note 2: In older versions of this protocol, an optional component of the lysis buffer was 10 mM vanadyl ribonucleoside complexes (VDR).⁷ VDR oxidizes very rapidly and therefore should be added to the lysis buffer just before use. It is usually prepared/supplied as a 200 mM stock solution and should be used at a working concentration of 5–20 mM. The use of VDR has fallen out of favor, however, because of its inhibitory effect on reverse transcriptase and other enzymes. Alternatively, RNasin® (Promega) can be added to the lysis buffer (250–1000 U/ml) to inhibit RNase activity. Beyond this point in the protocol, the tube containing the RNA should be maintained on ice unless otherwise indicated.

5. Pellet the nuclei and other cellular debris by differential centrifugation for 5 min at 4°. Centrifuge at 5000 \times g in a microfuge, or at a suitable g-force if another type of tube is used.

Note 1: Although nuclei can be efficiently pelleted with as little as 500 \times g, the added force is useful for the more complete removal of large organelles and the tight packing of the nuclei. This centrifugation will result in the removal of genomic DNA as well as the nuclear hnRNA species and will facilitate the analysis of extranuclear RNA.

Note 2: If the nuclei are to be used in a later application, such as the isolation of hnRNA or genomic DNA, then collect the nuclei at a maximum of 500 \times g, and extend the centrifugation time as need. Additional details may be found in Chapter 16.

6. Transfer the supernatant (cytosol) to a fresh, RNase-free tube and add EDTA (pH 8.0) to a final concentration of 10 mM.

Note: Na₂-EDTA, routinely prepared as a 500 mM stock solution, is added to prevent the magnesium-mediated aggregation of nucleic acids with proteins and with each other. Adding EDTA to the lysate prior to this step would have caused nuclear rupture and contamination of cytoplasmic RNA with nuclear transcripts as well as DNA.

7. Add an equal volume of the organic phase (lower phase; yellow) of the extracting buffer to the aqueous cytosol (lysate). Using extreme care, mix thoroughly by inversion for 15 s or longer. Incubate the sample at 55° for 5 min.

Note 1: If the prep is being done in a microfuge tube, hold the tube between the thumb and index finger and invert the tube several times. This technique will give the most complete intermixing of the phases, the most complete denaturation of proteins, and minimize the likelihood of accidental spillage.

Note 2: The heating of the sample will facilitate more thorough removal of proteins. Some variations of this procedure direct the investigator to do a series of organic

⁶Autoclave this lysis buffer *before* adding the NP-40 needed to complete this reagent.

⁷See Chapter 7 for discussion of advantages and disadvantages of VDR, and for a synthesis protocol.

extractions without heating the sample. The approach recommended here reduces the volume of organic waste.

Note 3: Upon contact with VDR, the initially yellow 8-hydroxyquinoline becomes dark green-black. This is because 8-hydroxyquinoline will chelate heavy metals (vanadium). In the absence of 8-hydroxyquinoline (or VDR), no color change is expected.

8. Incubate the sample on ice or in an ice water bath for 5 min.

9. Centrifuge the sample at $500\text{--}2000 \times g$ at 4° or at top speed in a bench-top microfuge to separate the phases.

Note: The time required to complete the separation will be a direct function of the volume and the size of the tube. A maximum of 5 min is usually adequate. Although a refrigerated microfuge is useful, it is not required.

10. Carefully transfer the upper aqueous phase to a fresh, RNase-free tube. Use great care to avoid disturbing the protein interface which appears as a thin white band between the upper and lower phases. It is much better to leave a few microliters of aqueous material behind than to risk transfer of any material at the interface.

11. Add a fresh aliquot (equal volume) of extracting buffer to the aqueous phase and mix carefully and thoroughly by inversion. Centrifuge at $500\text{--}2000 \times g$ at 4° or at top speed in a bench-top microfuge to separate the phases. It is not necessary to heat and cool the sample again, as described in steps 6–7.

12. Carefully transfer the upper aqueous phase to a fresh RNase-free tube. If a protein interface is apparent or if the aqueous phase remains cloudy, repeat the organic extraction described in Step 10 until no protein appears at the interface and the aqueous (upper) phase is clear.

Note: Additional phenol extractions are usually necessary when extracting a large number of cells in a relatively small volume. Using $100 \mu\text{l}$ lysis buffer per 10^6 cells at the beginning of this protocol usually precludes this difficulty.

13. Extract the aqueous material one last time with an equal volume of chloroform or chloroform:isoamyl alcohol (24:1) and mix carefully and thoroughly by inversion. Centrifuge for 30 s to separate the phases.

Note: Because phenol is very soluble in chloroform, a final chloroform extraction will remove any traces of phenol from the RNA-containing aqueous phase. This should be a standard technique whenever phenol is involved. Remember that even trace amounts of phenol will oxidize into quinones which will compromise the quality of a nucleic acid sample.

14. Transfer the aqueous phase (upper) to a fresh, RNase-free tube and add 0.1 volume 3 M sodium acetate, pH 5.2 and 2.5 volumes ice cold 95% ethanol. Mix thoroughly by inversion.

15. Precipitate the RNA overnight at -20° .

Note: The addition of salt and alcohol is the most versatile method for the precipitation of nucleic acids. A complex that has reduced solubility in high concentrations of alcohol forms consisting of the sample and the salt. The precipitate can then be collected by centrifugation, briefly dried, and resuspended in an appropriate buffer. To ensure complete RNA recovery, samples should be precipitated overnight at -20° , as the recovery of RNA from a biological source is not always predictable.

16. Collect the precipitated RNA by centrifugation at $9000 \times g$ at 4° for 30 min if using Corex glass or at $12,000 \times g$ in a microfuge for 10 min at 4° (preferred) or at room temperature.

17. Carefully decant the ethanol supernatant. Excess coprecipitated salt should be removed by washing the RNA pellet 2–3 times with 70–75% ethanol, into which

the salt will dissolve (sodium acetate is particularly soluble in 70–80% ethanol). This concentration of ethanol is not sufficiently aqueous, however, to redissolve precipitated nucleic acids.

18. Briefly air dry the tube(s). If desired, a final wash with 95% ethanol will facilitate the drying of the sample. The RNA is best stored at -80° as an ethanol precipitate or, following quantification, stored in suitable aliquots at -80° so as to avoid repeated freezing and thawing. Depending on the cell type, typical yields range from 75–100 μg cytoplasmic RNA per 10^7 cells.

Chaotropic lysis buffers

There is probably no better way to deal with seemingly recalcitrant RNases than to disrupt cells in guanidinium lysis buffer (Chirgwin *et al.*, 1979). Guanidinium buffers efficiently denature and solubilize proteins, including RNase, which inhibits their degradative prowess. This effect is enhanced by the inclusion of the ionic detergent sarkosyl. It is not necessary to add additional RNase inhibitors to such lysis buffers, and RNA isolation procedures conducted under these harsh conditions can be performed with room temperature reagents. Because these buffers are so chaotropic (biologically disruptive), organelle lysis also accompanies disruption of the plasma membrane. This occurrence liberates heterogeneous nuclear RNA (hnRNA) and genomic DNA from the nucleus as well as mitochondrial DNA and RNA, all of which copurify with the cytoplasmic RNA. Therefore, further purification is required to remove DNA from the sample. Formerly, the most prevalent of these methods was a technique known as isopycnic centrifugation (see Cooper, 1977; Castora and Greene, 2009 for background information) through a cesium chloride (CsCl) gradient (Glišin, 1974; Ullrich *et al.*, 1977) or a cesium trifluoroacetate (CsTFA) gradient (Carter *et al.*, 1983; Zarlenga and Gamble, 1987). Isopycnic separation is possible because of the differing buoyant densities of DNA (1.5–1.7 g/ml) and RNA (1.8–2.0 g/ml). For the convenience of the reader, a review of the fundamentals of centrifugation as a tool in molecular biology is presented in Appendix O.

Subsequently, the differential partitioning of DNA, RNA, and protein by guanidinium–acid–phenol extraction was described (Chomczynski and Sacchi, 1987) and is now one of the most commonly used techniques for rapid purification of RNA. In this approach, the salient chemical differences between RNA, protein, and DNA are exploited by creating an acidic pH environment and judiciously blending organic solvents. As a result of the ease of this approach, numerous products and methodologies have been developed for the rapid, efficient purification of both RNA and DNA (and protein) from the same biological source (Majumdar *et al.*, 1991; Chomczynski, 1993). Currently, the use of glass fiber filters (a.k.a. silica binding technology) is the most popular method for small-scale RNA isolation and is used in conjunction with guanidinium-based cell lysis. Some of these procedures, and others, are described here.

The principal drawback of all of these chaotropic procedures is that the investigator is not be able to discriminate between cytoplasmic and nuclear

RNA, as there is no method for separating hnRNA from spliced messenger RNA (mRNA) once these two populations have been mixed (size fractionation may result in a partial separation, but is not at all definitive).

It is unfortunate that many seasoned investigators begin to show signs of sloppiness with respect to keeping buffers and equipment nuclease free when routinely working with guanidinium buffers. Although it is true that RNA is safe from nuclease degradation in the presence of these agents, purified RNA is once again susceptible to nuclease degradation when the denaturants have been removed. Therefore, it is necessary to guard against the RNase peril consistently, according to the guidelines described in Chapter 7.

Isolation of RNA with guanidinium buffers

RNA lysis buffers that contain guanidinium thiocyanate or guanidinium-HCl reproducibly yield very high-quality RNA samples. This is true because of the extremely chaotropic nature that these chemicals exhibit; they are among the most effective protein denaturants (Cox, 1968; Nozaki and Tanford, 1970; Gordon, 1972). This attribute supports the inclusion of guanidinium as a dependable method when planning an RNA extraction from whole tissue samples, particularly those enriched in RNase. The nuclear and organelle disruption that accompanies plasma membrane solubilization liberates nuclear RNA and genomic DNA, both of which copurify with cytoplasmic RNA species. The efficiency of protein denaturation, including disruption of RNases, may be enhanced by the inclusion of β -mercaptoethanol (β -ME). This reducing agent acts to break intramolecular protein disulfide bonds. The addition of exogenous inhibitors of RNase is not required. Another common reducing reagent, dithiothreitol (DTT) should be avoided in this particular application because it is chemically reactive with guanidinium.

The use of guanidinium-based lysis buffers mandates a procedure for the partitioning of RNA, DNA, and protein in the resulting lysate. There are currently three basic approaches for accomplishing this task: examples of each are presented here. These approaches involve whole cell lysis followed by isopycnic ultracentrifugation, acid-phenol extraction, or binding to glass fiber filters (Fig. 2.3). Although the RNA is most often precipitated directly from a guanidinium-containing lysate, ultracentrifugation through CsCl (or CsTFA), though cumbersome, is useful when ultrapure RNA is required. Running gradients is generally a time-consuming procedure that is usually no longer required for mainstream molecular biology applications.

CAUTION: *If you are not familiar with the proper operation of an ultracentrifuge or the procedure for derating a rotor be sure to get proper technical assistance. If not used properly, a rotor can be lethal.*

In contrast to the gentle lysis methods, which facilitate the isolation of total cytoplasmic RNA, procedures involving the use of guanidinium-containing buffers are used in the isolation of total cellular RNA (nuclear and cytoplasmic

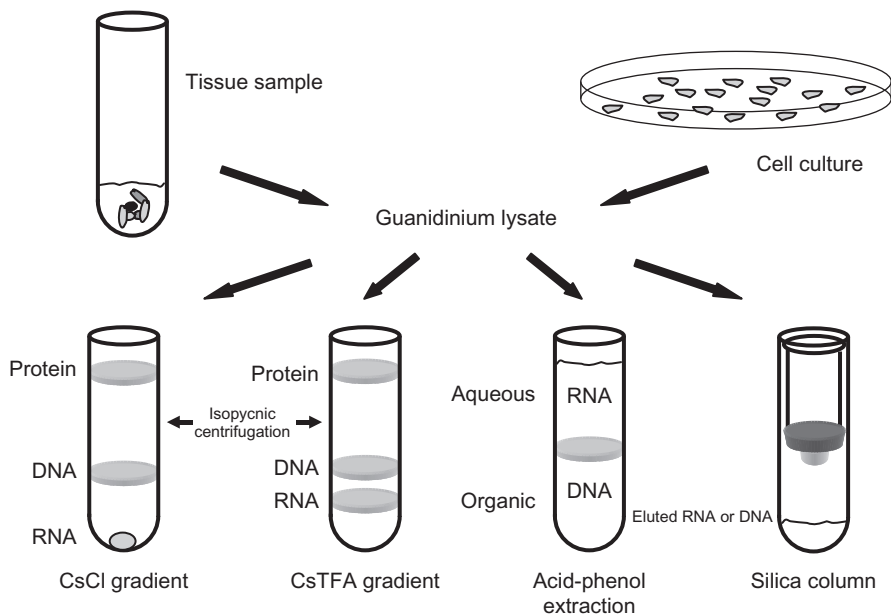


Figure 2.3 Methods for fractionation of total cellular RNA following cellular disruption with guanidinium buffer: CsCl or CsTFA ultracentrifugation, acid-phenol extraction, or binding to a glass fiber filter.

species). The RNA size distributions from total cytoplasmic and total cellular isolation procedures are compared in Fig. 2.4.

Guanidinium–acid–phenol extraction techniques

The highest quality RNA indisputably results from the extraction of RNA mediated by chaotropic lysis buffers, and guanidinium-containing buffers are among the most effective. As described above, these isolation procedures have the advantage of disrupting cells grown in culture or whole tissue samples rapidly and completely, simultaneously inactivating RNase activity, even when it is present in great abundance. The original procedures are labor-intensive because of the inclusion of an ultracentrifugation step to separate RNA from DNA based on differences in buoyant density. In contrast, Chomczynski and Sacchi (1987) describe the isolation and purification of undegraded RNA by treatment of cells with guanidinium thiocyanate-containing lysis buffers, but without the need for subsequent CsCl ultracentrifugation of the sample. In this and related procedures, the RNA is isolated in a very short time by extraction of a guanidinium cell or tissue lysate with an acidic phenol solution; chloroform is added to facilitate partitioning of the aqueous and organic material. Although such extraction buffers are easily prepared in the molecular biology laboratory

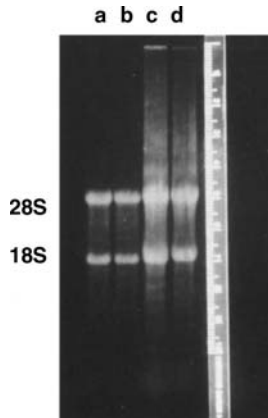


Figure 2.4 Comparison of size distribution of total cellular RNA and total cytoplasmic RNA. Lanes a and b: 20 μ g total cytoplasmic RNA isolated with NP-40 lysis buffer and subsequent deproteinization with phenol:chloroform, as described. Lanes c and d: 25 μ g of total cellular RNA isolated by guanidinium-acid-phenol extraction. The higher molecular weight species (hnRNA) are clearly visible above the 28S rRNA in lanes c and d. The obvious fluorescence in the wells of lanes c and d suggests a small amount of contaminating DNA. Fluorescent ruler courtesy of Diversified Biotech, Inc.

by mixing water-saturated phenol with an acidic solution of sodium acetate, pre-mixed monophasic formulations of phenol and guanidinium thiocyanate (e.g., TRIzol[®] and TRI Reagent[®]) are readily available from a number of vendors. These commercial reagents represent improvements to the original published protocol. Upon phase separation, RNA is retained in the aqueous phase, while DNA and proteins partition into the organic phase; a protein interface is generally not observed. RNA is then recovered by precipitation with isopropanol and collected by centrifugation. In these procedures, RNA can be efficiently isolated from as little as 1 mg of tissue or 10^6 cells, usually in less than 1 h.

Protocol: guanidinium–acid–phenol extraction

1. Wear gloves!
2. Collect harvested cells by centrifugation and resuspend cell pellets in 100 μ l of the following formulation (solution D) per 10^6 cells:
 - 4 M guanidinium thiocyanate
 - 25 mM sodium citrate, pH 7.0
 - 0.5% sarcosyl
 - 100 mM 2-mercaptoethanol (β -ME)

Micropipette up-and-down, or gently vortex in order to break up the cell pellet.

Note: Cells may be lysed directly by the addition of 1.0 to 1.5 ml of solution D per 100 mm tissue culture dish. Lysis in the culture dishes and flasks usually mandates the use of larger reagent volumes, and it may prove difficult to recover all of the

lysate due to its tremendous viscosity. In such an event, a sterile cell scraper (e.g. Corning item number 3010) may be helpful for the mechanical harvesting of cells from culture dishes or flasks. Scaled down, the entire procedure can be performed in a 2.2 ml microfuge tube.

3. Transfer the lysate to a polypropylene tube. For each 1 ml of solution D lysis buffer used in step 2, add:
 - 0.1 ml 2 M sodium acetate, pH 5.2
 - 1.0 ml water-saturated phenol (molecular biology grade)
 - 0.2 ml chloroform:isoamyl alcohol (49:1)

Cap tube and mix carefully and thoroughly by inversion following the addition of each reagent and invert vigorously for an additional 30 s after all reagents have been added.

4. Cool sample on ice for a minimum of 15 min; centrifuge at 4° to separate the phases.
5. Transfer aqueous (upper) phase containing the RNA to a fresh tube and mix with 0.75 volume of ice-cold isopropanol. Store at -20° for at least 1 h to precipitate RNA.

Note: For larger scale preparations, Corex glass tubes, placed in the correct adapters, can be used for precipitation and recovery of RNA.

6. Collect precipitate by centrifugation at 10,000 × g for 20 min at 4°. Carefully decant and discard supernatant. **CAUTION:** Do not exceed the recommended maximum g-force for any of the tubes used in this protocol.
7. Completely dissolve RNA pellet in 300 μl of solution D (see step 2) and then transfer to a RNase-free 1.7 ml microfuge tube.
8. Reprecipitate the RNA by the addition of 0.75 volume of ice-cold isopropanol and store at -20° for 1 h.
9. Collect precipitate at 12,000 × g in a microcentrifuge for 10 min at 4°. Carefully decant and discard supernatant.
10. Wash pellet 3 to 4 times with 500 μl 70% ethanol per wash. If the RNA does not dislodge during these washes, there is no need to re-centrifuge. Allow tubes to air-dry to remove residual ethanol.

Note: If desired, a final wash with 500 μl of 95% ethanol will accelerate the drying process.

11. Redissolve RNA in the smallest possible volume of TE buffer or nuclease-free H₂O. Incubation at 65° for 10 min may facilitate solubilization, though this is unnecessary if the RNA did not dry out completely following the ethanol washes. Store the RNA as an ethanol precipitate until it is to be used. Following determination of concentration, store the RNA in suitable aliquots at -80°. Avoid repeated freezing and thawing.

Density gradient centrifugation

The original protocols and several current procedures for the isolation of RNA from biological sources enriched in RNase described cell and tissue disruption with guanidinium buffer. The resulting intermixing of subcellular components mandates the separation of these biochemical macromolecules from each other, with particular regard to the removal of DNA from RNA preparations. Subtle, though measurable, differences in the densities of DNA, RNA, and protein (Table 2.3) allow fractionation by banding or pelleting them in a density gradient.

Density gradient centrifugation, known more properly as isopycnic centrifugation, is a technique in which macromolecules move through a density gradient until they find a density equal to their own. The classical gradients for this type of separation are CsCl, CsTFA, and cesium sulfate (Cs₂SO₄)⁸. Macromolecules therefore accumulate at this position in the gradient, floating there until the end of the centrifugation run. In some cases, RNA is of greater density than any position in the gradient and accumulates as a pellet at the bottom of the tube, whereas gradients of greater maximum density may permit the banding of RNA, much as plasmid DNA is banded for purification.

Cesium chloride

Cesium chloride (CsCl) is a dense salt that exhibits the ability to form a linear gradient when ultracentrifugation g-forces are applied to a homogeneous suspension (e.g., cell lysate). There is no need to pre-form the gradient. CsCl chloride gradients are generally steep, and the maximum density within the gradient usually exceeds all but the densest material to be sedimented, meaning that higher density molecules such as RNA will collect as a pellet at the bottom of the tube. CsCl gradients are subjected to ultracentrifugation g-forces for a period sufficiently long to complete the migration of all macromolecules into their equilibration positions. The actual time required is a function of the rotor size; traditional RNA preps in floor model ultracentrifuges require overnight centrifugation, while only a few hours are needed to accomplish the same type of separation in table-top or “micro-ultracentrifuges,” for example, Sorvall Discovery™ M150 SE (Fig. 2.5).

Table 2.3 Typical Buoyant Density Ranges of Macromolecules Resolved by Isopycnic Centrifugation

Molecule	Density
Protein	1.2–1.5 g/ml
DNA	1.5–1.7 g/ml
RNA	1.7–2.0 g/ml

Protocol: Cesium chloride (CsCl) gradients

The RNA isolation protocol presented here is a modification of the procedures described by Glišin *et al.*, (1974), Ullrich *et al.*, (1977), and Chirgwin *et al.*, (1979).

⁸Other materials such as glycerol, ficoll, metrizamide, and sucrose were also used at one time primarily because they are readily available at low cost and high purity. Nucleic acid resolution based on differences in density is best performed using one of the cesium salts.



Figure 2.5 Sorvall Discovery™ M150 Micro-ultracentrifuge. Smaller radius rotors permit higher g-forces which, in turn, permit isopycnic preparation of RNA in 3 hours or less. Photo courtesy of Kendro Laboratory Products.

1. Wear gloves!
- 2a For tissue culture: collect harvested cells by centrifugation at $500 \times g$ for 5 min at 4° (if possible). Decant as much of the supernatant as possible. Wash cells once with ice-cold PBS and then decant as much PBS as possible.
- 2b For tissue: *rapidly* harvest the tissue and mince/dice in ice cold saline solution. Quickly transfer tissue into a suitable homogenization vessel that already contains an aliquot of lysis buffer (step 3). If tissue was flash-frozen in liquid nitrogen, it may be helpful to shatter it (while wrapped in aluminum foil) by impact upon the lab bench. In so doing, the sample can be quickly transferred and allowed to thaw in the lysis buffer.

Note: See Chapter 3 for information on the subtleties of working with tissue.

3. Add 5 volumes (the size of the pellet) of a solution consisting of:
 - 6 M guanidinium hydrochloride (or 4 M guanidinium thiocyanate)
 - 5 mM sodium citrate (pH 7.0)
 - 0.5% sarcosyl
 - 10 mM β -ME (add just prior to use).

Gentle vortexing may facilitate dissociation of the pellet.

Note 1: Guanidinium thiocyanate solution may be stored at 4° for several days, prior to the addition of β -ME. Immediately prior to use, warm a suitable aliquot of the guanidinium thiocyanate solution to 37°, cool to room temperature, and then add β -ME to a final concentration of 200mM. If desired, the solution can be clarified by filtration through a 0.2 μ m filter.

Note 2: Be sure to check for guanidinium buffer- and β -ME incompatibility with ultracentrifuge tubes.

4. Disrupt the cells or tissue at room temperature by either manual homogenization (repeatedly drawing the lysate through a 19-gauge needle; several strokes using a Dounce homogenizer) or mechanical homogenization (e.g., Polytron).

Note 1: Shearing forces physically break up genomic DNA so as to preclude the formation of an impenetrable web of high-molecular-weight DNA that would otherwise impede the sedimentation of RNA.

Note 2: For optimal recovery of high-quality RNA, avoid generating excess heat or foam during this step.

5. If desired, add 1 g of solid CsCl for every 2.5 ml of lysate. This is required only when other molecules are to be copurified from the same gradient along with the RNA.
6. Pipette a cushion of 5.7M CsCl (dissolved in 100mM EDTA, pH 7.5) into an ultracentrifuge tube, typically cellulose nitrate or polyallomer, compatible with the swinging bucket rotor to be used. The 5.7M CsCl cushion should occupy about 25% of the total volume of the tube.
7. Gently layer the lysate on top of the CsCl cushion until the ultracentrifuge tube is filled. Be sure to weigh each centrifuge tube along with its swinging bucket and corresponding top. Each sample (bucket, screw top, centrifuge tube containing the sample) should weigh within 0.05 g of the other samples to be loaded onto the rotor. Follow the manufacturer's instructions for sealing the tubes, and be sure to weigh each sealed sample again before loading the rotor. Failure to balance precisely the rotor buckets according to the specifications of the centrifuge manufacturer will have catastrophic consequences.
8. Centrifuge the preparation in a swinging bucket rotor. Floor model ultracentrifuges generally require an overnight run, while micro-ultracentrifuge models require only 3–4 h. Each rotor instruction guide will delineate the correct relative centrifugal force (RCF) needed to purify RNA safely using this type of centrifugation.

Note: Perform this type of centrifugation at ambient temperature. Cooling the centrifuge chamber at any time during the run will result in precipitation of the CsCl and ruination of the gradient.

9. At the conclusion of the run, carefully remove the supernatant by aspiration (sterile Pasteur pipettes, or sterile 5 ml pipettes with the tip removed, work well). Use great care not to disrupt the pellet at the bottom of the tube. Many investigators prefer to remove 80% of the gradient by pipetting, and the remainder by decanting. After the supernatant has been removed, it is useful to cut off the top portion of the ultracentrifuge tube just above the RNA pellet. In this lab, a sterile razor blade is clamped in a hemostat, briefly heated with a Bunsen burner, and used to slice

through the emptied centrifuge tube. In so doing, the translucent RNA pellet will be easily accessible in what looks like a giant contact lens-shaped cradle (Fig. 2.6).

Note: This technique will help prevent contamination of the RNA pellet with protein and/or DNA that may be clinging to the side of the ultracentrifuge tube, even after removal of the gradient.

10. Dissolve the RNA pellet completely in the smallest possible volume of:

10 mM Tris, pH 7.4

5 mM EDTA

0.1% SDS (optional)

and then transfer the sample to a sterile microfuge tube. For large-scale RNA extractions, transfer the RNA suspension to a conical 15 ml polypropylene tube.

Note: The RNA pellet is often difficult to dissolve in aqueous buffer after this type of purification, and the larger the sample, the more difficulty will be encountered. Freezing the pellet at -20° and then thawing usually facilitates dissolving it and removal from the centrifuge tube. Limited vortexing may help.

11. Extract the sample once with 2 volumes of chloroform:isoamyl alcohol (4:1, v/v) or with 2 volumes of chloroform:butanol (4:1, v/v). Pulse centrifuge to separate the two phases. Transfer the upper aqueous phase to an autoclaved microfuge tube.

12. Re-extract the organic phase and interface with an equal volume of:

10 mM Tris, pH 7.4

5 mM EDTA

0.1% SDS (optional)

and combine the two aqueous phases.

Note: These chloroform extractions are needed to remove residual contaminants from the RNA preparation.

13. Precipitate the RNA for several hours at -20°C , with the addition of 0.1 volume 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice-cold 95% ethanol. Depending on the volumes involved, it may be necessary to split the sample into two or more tubes or a 15 ml Corex glass tube.

14. If the RNA was precipitated in a microfuge tube, collect the precipitate at $12,000 \times g$ for 10 min, at 4° if possible. If precipitation was carried out in a Corex glass tube, centrifuge in a suitable adapter for at least 20 min at $9000 \times g$. **CAUTION:** Be certain

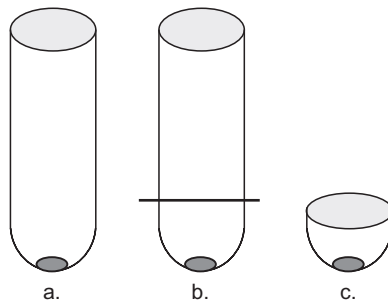


Figure 2.6 Method for recovery of RNA following isopycnic centrifugation. (a) Appearance of RNA pellet after removal of gradient. (b) Ultracentrifuge tube is cut just above the RNA. (c) RNA rests in contact lens-shaped cradle. This approach favors recovery of RNA free of DNA and protein contamination from the inside walls of the tube.

not to exceed the maximum g-force rating for the rotor or the type of tube in which the RNA was precipitated.

15. Carefully decant the supernatant and wash the pellet at least once with 70% ethanol. Decant the 70% ethanol wash and, if desired, wash the sample once with 95% ethanol to accelerate the drying of the sample. Do not allow the pellet to dry out completely.
16. The RNA is best stored at -80° as an ethanol precipitate or, following quantification, in suitable aqueous aliquots at -80° to avoid repeated freezing and thawing. Depending on the cell type, typical yields range from 50–75 μg of very high-quality cellular RNA per 10^6 cells.

Cesium trifluoroacetate (CsTFA)

CsTFA (GE Healthcare Life Sciences) is an excellent CsCl alternative for density gradient separation of RNA and DNA by isopycnic centrifugation. At ultracentrifugation RCFs, CsTFA will likewise self-form a gradient (Fig. 2.7). The trifluoroacetate anion gives CsTFA properties that result in very high-quality nucleic acid preparations of extremely high quality, including (1) more ready dissociation of proteins from nucleic acids; (2) the ability to band, rather than pellet RNA, because of high solution density; (3) simultaneous banding and recovery of DNA, RNA, and protein from the same CsTFA gradient (Zarlenga and Gamble, 1987); (4) improved yields of pure RNA due to the salting-in effect of the trifluoroacetate ions, in contrast to the salting-out effect of the chloride and sulfate anions which causes protein precipitation and concomitant loss of nucleic acids during isolation; (5) solubility in polar solvents, such that nucleic acids may be recovered directly from the gradient by precipitation with ethanol and without prior dialysis or extraction to remove the salt; and (6) excellent inhibition of RNase.

Disadvantages

Among the disadvantages of such density gradients are the cost of the density gradient material (e.g., CsCl, CsTFA), the requirement for a very expensive ultracentrifuge, the relatively long period of centrifugation, the labor-intensive clean up of gradient-purified RNA, and the limited ability to prep more than a few samples at a time. Thus, isopycnic centrifugation has become something of a rarity in the isolation of RNA, in favor of newer reagents and kits. In some applications, there remains the need to resolve RNA in a density gradient and for this reason, the CsCl protocol (above) and the CsTFA protocol are included here.

Protocol: cesium trifluoroacetate (CsTFA) gradients

This protocol is a modification of the procedure of Okayama *et al.* (1987). As presented, the RNA will form a pellet at the bottom of the tube, rather than banding within the gradient. It is recommended for the preparation of total RNA from a variety of tissues and can be modified to accommodate as little as 25 mg or as much as 6 g of starting material. Do not use cellulose nitrate tubes when running a CsTFA gradient.

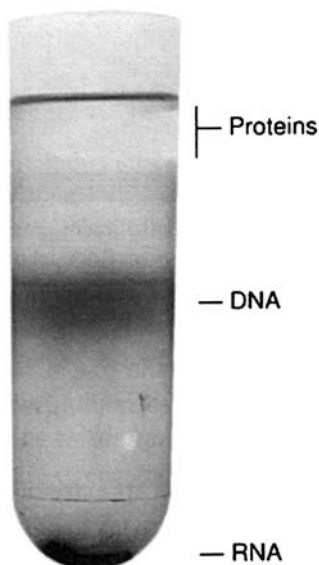


Figure 2.7 CsTFA gradient. Isopycnic gradient separation of calf liver macromolecules extracted with guanidinium thiocyanate. A guanidinium thiocyanate extract of calf liver (18 ml) was layered on top of a 19 ml cushion of CsTFA ($\rho = 1.51$ g/ml) and centrifuged at 125,000 g for 16 h. Approximately 3.5 mg of total RNA was pelleted at the bottom of the tube ($\rho \approx 1.62$ – 1.9 g/ml). This RNA contained less than 0.4% DNA as measured with a DNA probe. DNA forms a band near the original interface of the cushion and the sample solutions ($\rho \approx 1.6$ g/ml). Protein in the sample collects in the upper portion of the gradient ($\rho \approx 1.2$ – 1.5 g/ml). Courtesy of Amersham Biosciences.

Advance Preparation of CsTFA

CsTFA (GE Healthcare Life Sciences No. 17-0847-02) is usually provided at a density of 2.0 ± 0.05 g/ml, and prepared as a working solution with a density of 1.51 ± 0.01 g/ml, in 100 mM EDTA, pH 7.0. To prepare 100 ml of this solution, mix $51/(\rho-1)$ ml of CsTFA (where ρ = density of product as received) with 40 ml of a solution of 100 mM Tris (pH 8.0), 100 mM KCl, 1 mM EDTA, and $60 - [51/(\rho-1)]$ ml of sterile distilled water. The density of the working solution of CsTFA may be altered to fit a particular experimental protocol and additional formulae for preparing precisely defined densities of CsTFA are provided in the product insert that accompanies this product.

1. Wear gloves!
- 2a. For tissue culture: Collect harvested cells by centrifugation at $500 \times g$ for 5 min. Decant as much of the supernatant as possible. Wash cells once with PBS (per liter: 8.0 g NaCl; 0.2 g KCl, 1.44 g Na_2HPO_4 , 0.24 g KH_2PO_4). Decant as much PBS as possible. If the RNase is likely to be a major challenge, perform centrifugations at 4° , and wash cells with ice-cold PBS.
- 2b. For tissue: *rapidly* harvest the tissue and mince in ice-cold saline solution. Quickly transfer tissue into a suitable homogenization vessel that already contains an aliquot

of lysis buffer (step 3). If tissue was flash-frozen in liquid nitrogen, it may be helpful to shatter it (while wrapped in aluminum foil) by impact upon the lab bench. In so doing, the sample can be quickly transferred and allowed to thaw in the lysis buffer.

Note: See Chapter 3 for the subtleties of working with tissue.

3. For each gram of tissue or cell pellet, add 18 ml of 5.5 M guanidinium thiocyanate solution (5.5 M guanidinium thiocyanate, 25 mM sodium citrate, 0.5% sarcosyl; adjust pH to 7.0 with NaOH; add 200 mM β -ME just prior to use).

Note 1: Guanidinium thiocyanate solution may be stored at 4° for several days, prior to the addition of β -ME. Immediately prior to use, warm a suitable aliquot of the guanidinium thiocyanate solution to 37°, cool to room temperature, and then add β -ME to a final concentration of 200 mM. If desired, the solution can be clarified by filtration through a 0.2 μ m filter.

Note 2: Be sure to check for guanidinium buffer-, CsTFA, and β -ME incompatibility with ultracentrifuge tubes.

4. Disrupt the cells or tissue at room temperature by either manual homogenization (repeated shearing through a 19-gauge needle; several strokes using a Dounce homogenizer) or mechanical (e.g., Polytron) homogenization.

Note 1: Shearing forces physically break up genomic DNA so as to preclude the formation of an impenetrable web of high-molecular-weight DNA that would otherwise impede the sedimentation of RNA.

Note 2: For optimal recovery of high-quality RNA, avoid generating excess heat or foam during this step.

5. Carefully transfer the lysate to disposable conical centrifuge tubes (e.g., 50 ml size). Centrifuge at $1500 \times g$ for 5 min to sediment insoluble materials.
6. Without disturbing the pelleted material, transfer the clarified lysate to a fresh tube.
7. If the lysate still appears viscous, shear the DNA by once again drawing it through a 19-gauge needle repeatedly until the viscosity decreases. Between 10 and 12 passes into and out of the syringe is usually adequate.
8. Centrifuge at $5000 \times g$ for 20 min at 15° to remove the cell debris. Decant and retain the supernatant, taking great care not to dislodge the pelleted debris. **CAUTION:** Do not exceed the maximum g-force recommended for any of the centrifuge tubes used in this protocol.
9. If necessary, adjust the volume of the supernatant with 5.5 M guanidinium thiocyanate solution (see step 3).
10. Select a rotor suitable for the number and volume of samples. DO NOT use cellulose nitrate tubes with CsTFA. Pipette cushions of CsTFA into each tube, and then overlay aliquots of the sample onto these cushions, filling each tube to the very top. The CsTFA cushion should occupy about 25% of the total volume of the tube.
11. Be sure to weigh each centrifuge tube along with its swinging bucket and corresponding top. Each sample (bucket, screw top, and centrifuge tube containing the sample) should weigh *exactly* the same as the other samples to be loaded onto the rotor, within 0.05 g of each other. Follow the manufacturer's instructions for sealing the tubes, and be sure to weigh each sealed sample again before loading the rotor. Failure to balance precisely the rotor buckets according to the specifications of the centrifuge manufacturer will have catastrophic consequences.
12. Centrifuge at $125,000 \times g$ in a swinging bucket rotor for 16–20 h at 15°. Under these conditions RNA will pellet at the bottom of the tube and DNA will collect in a band in the lower third of the gradient.

13. After centrifugation, carefully aspirate off most of the liquid in each tube, stopping as soon as the DNA band has been removed. If the DNA band is not visible, discontinue aspiration within 1 cm from the bottom of the tube. Decant the remaining liquid, taking care not to disturb the RNA pellet. Invert the tubes and allow them to drain onto a stack of 3–4 Kimwipes for 5 min.
14. Carefully cut off the bottom of each tube with a sterile razor blade or scalpel, as shown in Fig. 2.6. Leave a sufficient portion of the sides of the tube to form a small cup.
15. Dissolve RNA pellets directly in the tube bottoms using two aliquots of an appropriate buffer (nuclease-free H₂O or TE, pH 7.4 (10 mM Tris-Cl, pH 7.4; 0.1 mM EDTA).
Note: At least two small aliquots (100–200 μl) of buffer are recommended to recover all of the RNA sample from the tubes. Repeated pipetting may be necessary for complete recovery. As always keep the sample as concentrated as possible so that all manipulations, including reprecipitation if necessary, can be carried out in a microfuge tube.
16. Combine both aliquots of RNA and vortex. Heat tubes to 65° for 10 min and vortex again.
17. Spin tubes briefly and at low speed to remove any insoluble material.
18. Determine RNA concentration, purity, and yield (Chapter 6). If desired, RNA can be precipitated to concentrate the RNA sample and/or change buffers.
19. Store RNA in suitable aliquots at –80° until they are ready for use.

Simultaneous isolation of RNA and DNA

Many noteworthy procedures representing varying degrees of complexity have been described for the isolation of RNA from cells and tissues. Most biotech companies now offer devices, reagents, solutions, or resins for the “painless” isolation of RNA from numerous biological sources. These products vary widely in their simplicity, speed, cost, and effectiveness. Several procedures have also been described for the simultaneous recovery of RNA, DNA, and protein from the same biological source and some of these, too, are now available commercially.

The method presented here has the advantage of requiring no specialized equipment other than what is found even in the most modest of molecular biology labs. This protocol is a modification of the procedure of Majumdar *et al.* (1991) and is based on the selective partitioning of RNA and DNA by adjusting the pH of a phenolic lysate (Chomczynski and Sacchi, 1987; Chomczynski, 1993) but without the use of guanidinium reagents. Succinctly, the RNA is recovered in the aqueous phase by first extracting the sample at acidic pH; DNA in the sample is retained at the interface and in the organic phase. Following transfer of the RNA-containing aqueous phase to a fresh tube, the DNA is eluted from the organic phase by establishing an alkaline pH. This is known as a reverse extraction. In this procedure, the investigator is afforded the option of making the phenolic lysate acidic by the addition of sodium acetate, pH 4.9, or simply extracting the SDS/EDTA lysate (pH 8.7) with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), the overall pH of which is approximately 7.6.

Cells may be lysed directly in the tissue culture vessel or harvested first. Cell pellets should be gently dispersed prior to addition of the lysis buffer because shearing of the DNA is likely to occur when an aggressive attempt is made

to resuspend a cell pellet *following* addition of the lysis buffer. In this regard the investigator is strongly urged not to succumb to the temptation to vortex this or any lysate from which genomic DNA is to be purified, as the enormous shearing forces created will damage DNA in a very short time. The size distribution of the resulting fragments is likely to be unacceptably low.

Protocol: simultaneous isolation of RNA and DNA

1. Wear gloves!
2. In advance: Prepare phenol:chloroform:isoamyl alcohol extraction buffer as described in Fig. 2.2, with the exception that the phenol should be *water*-saturated 2–3 times, rather than being Tris-saturated.
- 3a. For tissue culture: Remove growth medium by aspiration and wash the monolayer twice with ice-cold PBS. If desired, it may be helpful to prepare the cell cultures for lysis by cooling them directly on a bed of ice for 15 min. Remove, by aspiration, as much PBS as possible. Add 2 ml of ice cold lysis buffer (10 mM EDTA, pH 8.0; 0.5% SDS) per 75 cm² flask containing up to 5×10^6 cells (for higher density cell cultures, use 100 μ l lysis buffer per cm²). Rock the flask gently from side to side to ensure lysis of all cells. Incubate on ice for 3 min. If desired, the lysate can be transferred to lyse cells in a second flask, followed by incubation for an additional 3 min. Gentle agitation throughout may help. Do not use an aliquot of lysis buffer for more than two tissue culture flasks; excessive DNA/RNA/protein concentrations in the lysate will compromise partitioning efficiency. Cells that were harvested from tissue culture and pelleted should be lysed using 200 μ l lysis buffer per 10^6 cells.

Note: Cooling reagents and cells on ice is an excellent method for controlling RNase activity. It is the responsibility of the investigator, however, to empirically determine whether a 15 min incubation on ice prior to lysis will change the cellular biochemistry enough to influence the outcome of the study.

- 3b. For tissue: mince freshly isolated tissue (up to 100 mg) on ice. For each milligram of tissue, add 250 μ l lysis buffer (10 mM EDTA, pH 8.0; 0.5% SDS) and gently Dounce homogenize at room temperature. Do not use a Polytron to homogenize, as this will cause unacceptable shearing of the genomic DNA to be recovered later in this protocol.

Note: See Chapter 3 for a discussion of the subtleties of working with tissue.

4. Transfer the lysate into a pre-chilled 15 ml polypropylene tube. When scaling down, a polypropylene microfuge tube may be satisfactory.

Note: Due to the viscous nature of liberated chromatin, a sterile cell scraper may be useful for the quantitative recovery of the lysate.

5. Add an equal volume of cold organic extraction buffer, prepared as described in step 2. Be sure to check for centrifuge tube compatibility with phenol and chloroform.
6. Carefully mix the tube(s) by gentle inversion several times. Do not vortex! The yield will improve dramatically if the tubes are maintained on ice at all times when not mixing.
7. Centrifuge samples in an appropriate rotor in order to separate the aqueous material from the organic material. **CAUTION:** *Be careful not to exceed the recommended rotor speed or the maximum g-force recommended for any of the tubes used in this protocol.*

Recovery of RNA

8. Transfer the aqueous (upper) phase to a fresh tube, pre-chilled and resting on ice. Take care to avoid disrupting the protein interface. Recovery of DNA, which must

- be eluted from the interface and organic phase, can be delayed until all steps pertaining to RNA recovery have been completed. Leave tubes on ice until then.
9. Extract the aqueous phase with a fresh, equal-volume aliquot of the cold organic mixture described in step 2.
 10. Centrifuge as described above. Carefully recover the aqueous phase and transfer to a nuclease-free Corex glass tube (or microfuge tube), taking care to avoid disrupting the interface or organic material.
 11. Add 500 μ l of ice-cold 1 M Tris-Cl, pH 8.0 and 200 μ l of 5 M NaCl for every 4 ml of RNA-containing aqueous material. Mix carefully and thoroughly. Add 2.5 vol of ice-cold 95% ethanol and mix thoroughly. Store at -20° for at least 1 h.
 12. Collect RNA precipitate by centrifugation at $9000 \times g$ for 10 min in a suitable tube. Centrifugation at 4° is preferred, though not essential. **CAUTION:** *Do not exceed the recommended maximum RCF for the centrifuge tube.*
 13. Carefully decant and discard the supernatant.
 14. Wash pellet 2–3 times with 70% ethanol. Centrifuge again to recover the precipitate. Carefully decant and discard supernatant.
 15. Allow sample to air-dry. If desired, a final wash with 95% ethanol may accelerate air drying.
 16. Dissolve the RNA pellet in 200 μ l of ice-cold TE buffer (10 mM Tris-Cl, pH 8.0; 0.1 mM EDTA). Incubate on ice for 1 h and then transfer to a fresh microfuge tube. A second aliquot of ice-cold TE buffer may be used to rinse the centrifuge tube and should be combined with the first aliquot.
 17. Add 6 μ l of 5 M NaCl and 800 μ l of ice-cold ethanol for every 300 μ l RNA solution. Incubate on ice for 5–10 min to precipitate the RNA. The sample can now be stored as an ethanol precipitate at -80° or used immediately.
 18. Collect RNA precipitate by centrifugation at $12,000 \times g$ in a microfuge for 8 min, preferably at 4° . Carefully decant the supernatant and dry to remove excess ethanol as described in steps 14–15.
 19. Resuspend the RNA pellet(s) in a minimum volume of the desired buffer (e.g., nuclease-free water) and allow 15 min on ice for the sample to dissolve. Depending on the degree to which the sample has dried, longer incubation on ice may be necessary to dissolve the sample completely. Recommended: incubate the sample with RNase-free DNase (Appendix F) to remove any contaminating DNA.
 20. Calculate the concentration, purity, and yield of the RNA. Store hydrated samples in suitable aliquots at -80° .

Recovery of DNA

21. Return to the tubes containing organic material saved from step 8. Remove any remaining aqueous material as completely as possible. Any remaining RNA will be hydrolyzed in the steps that follow. Alternatively, the purified DNA may be incubated with RNase (Appendix G) to remove all contaminating RNA at the end of the procedure.
22. To the remaining organic/interface material add an equal volume of 1 M Tris base solution. This solution should not be pH-adjusted (pH \approx 10.5). Mix thoroughly by gentle shaking. Do not vortex.

Note: This step is known as reverse extraction; by creating a strongly alkaline environment the DNA will be drawn into the aqueous phase.

23. Centrifuge samples at $2000 \times g$ for 15 min, at 4° if possible, in order to separate the aqueous material from the organic material. Carefully transfer the aqueous (upper) phase to a fresh tube.
 24. Repeat steps 21 and 22, but do not mix the two aqueous phases. Keep them in separate centrifuge tubes.
 25. To each Tris:DNA aqueous sample, add an equal volume of chloroform:isoamyl alcohol (24:1). Mix carefully and thoroughly. Pulse centrifuge samples for 30 s to separate the phases. Carefully transfer the aqueous (upper) phase from each tube to fresh tubes.
 26. Precipitate the DNA with the addition of 0.1 volume of 3M sodium acetate, pH 5.2, and 2.5 volumes of 95% ethanol. Mix thoroughly, first by gentle swirling and then by thorough inversion. Do not vortex.
- Note: Adding ethanol should cause the immediate precipitation of the genomic material, which, characteristically, should appear quite stringy. The DNA can be collected immediately by low-speed centrifugation ($500 \times g$) for 1 to 2 min. Alternatively, genomic DNA can be “fished” out of the ethanol solution using a silanized Pasteur pipette (Appendix I) such that there is minimal sticking of the DNA to the glass.*
27. Transfer the precipitated DNA to a fresh microfuge tube. Wash the DNA three times with $500\mu\text{l}$ aliquots of 70% ethanol to remove excess salt. A final wash with 95% ethanol will accelerate the drying process. Allow DNA to redissolve in a suitable buffer (e.g., TE buffer, pH 8.0) and then determine DNA concentration.

The word on kits

The investigator who prepares in-house reagents “from scratch” for RNA isolation is something of a rarity now because so many vendors are marketing products that purport to save both time and money. It is certainly true that kits can increase productivity, though it is essential for the investigator to ascertain that a kit which looks and sounds attractive (1) produces RNA that will be compatible with downstream applications; and (2) yields the population of RNA that is needed (e.g. total cellular, cytoplasmic, poly(A)⁺). Most importantly, don’t be afraid to pick up the telephone and call the local sales rep or even the national sales office to politely request an evaluation kit (read: free kit). If you are genuinely interested in a particular product, most vendors will be happy to provide you with a one-time sample, or at least proffer a substantial discount.

The more desirable kits minimize or altogether eliminate the use of the organic solvents phenol and chloroform, and do so mostly through the use of silica columns. If a kit or reagent does utilize either or both of these hazardous compounds, the isolation procedure has been optimized to require the smallest possible volumes. Be sure that residual organic materials, regardless of the volume, are disposed of in accordance with in-house policies.

Perhaps the newest and best innovation in “kit technology” over the past two to three years has been the development of products that are designed to

recover RNA molecules that are less than 200 bases, rather than discarding them as undesirables. The exigency of this technological improvement is apparent when one considers the tremendous importance of miRNA and other small molecules in the regulation of gene expression.

Silica technology

One of the important improvements in the area of nucleic acid isolation has been the development of silica filters that are small enough for use with a standard microcentrifuge. These filters consist of glass microfibers that are positioned in the bottom of small plastic insert that fits inside a standard 1.5 ml microcentrifuge tube (Fig. 2.8). Not only are these products used to purify RNA from biological sources, they can also be used to clean up RNA or DNA after various manipulations, such as restriction enzyme digestion, ligation reactions, cDNA synthesis, and PCR amplifications. Silica filter-based products are readily available from any number of vendors. In general, RNA (or DNA) is bound to silica in a high-salt, chaotropic environment produced by diluting a nucleic acid sample in guanidinium thiocyanate. Following a series of washes, the purified material is eluted from the matrix under very low salt conditions. Best of all, the nucleic acid purification and cleanup is performed in a remarkably short time.



Figure 2.8 Silica filter for nucleic acid isolation. Under chaotropic, high-salt conditions, nucleic acids bind to the silica insert, while lysis buffer components and other cellular macromolecules are washed away. The purified sample is then eluted under low-salt conditions. The most commonly used filter format used in the laboratory is small enough to fit into a standard 1.7 ml microcentrifuge tube (2.2 ml microcentrifuge tube shown). Courtesy of Sigma-Aldrich.

Isolation of cytoplasmic RNA on a silica column

Many kits have been developed for the isolation of DNA and RNA since all nucleic acids have a strong binding affinity for silica filters under high-salt (chaotropic) conditions. Thus, the preliminary handling of the biological material and the lysis buffer formulation determine whether DNA or RNA will be recovered at the end of the procedure. For example, cell lysis might be performed with a quasi-alkaline lysis buffer to which RNase may have been added, resulting in DNA isolation and RNA obliteration. Alternatively, incubation of an RNA prep with RNase-free DNase will usually render a very respectable RNA yield, assuming that standard RNase-free techniques were observed throughout. One might even consider gentle cell lysis with NP-40, to remove the nucleus after which the cytosol is processed as if working with a whole-cell suspension. This will render only cytoplasmic RNA, rather than cellular RNA, and has all of the advantages associated with silica column chromatography. This approach to nucleic acid isolation is very popular because it can be used with small amounts of tissue or cells, can be performed on the bench quickly, easily, and economically with standard equipment, and eliminates the need for any organic solvents. However, these systems often contain guanidinium buffers, which are hazardous, and should be handled according to the manufacturer's recommendations and material safety data sheet (MSDS).

Affinity matrices

In addition to the approaches described above for the isolation of total cellular RNA or total cytoplasmic RNA, there are products available which capture polyadenylated transcripts. For example, mRNA isolation kits feature tracts of oligo(dT) that have been covalently linked to a solid support, such that polyadenylated transcripts are captured through canonical base-pairing between the poly(A) tail and the oligo(dT) in a high-salt environment. Some kits feature columns that are pre-packed with oligo(dT) linked to cellulose, while in other cases the oligo(dT) is linked to polystyrene and/or latex beads. Affinity selection in this manner is similar to the magnetic separation methods described in Chapter 5, though the affinity separation systems described here are not based on magnetic separation. The benefit associated with affinity selection is enrichment of a nucleic acid sample in favor of mRNA by minimizing, if not altogether eliminating, carryover of rRNA and tRNA; enrichment in this manner may increase the sensitivity of an assay. An older type of affinity selection in which poly(A)⁺ mRNA was affinity-captured using a column packed with poly(U) linked to sepharose beads (Lindberg and Persson, 1974) is no longer favored because of the relatively poor binding capacity of poly(U) matrices and the fact that quantitative recovery of RNA from the a poly(U) matrix usually required formamide-based elution buffers.

Yet another variant of the affinity matrix approach is designed to study nucleic acid protein interactions by passing a heterogeneous protein mixture

over a column packed with either RNA or DNA oligonucleotides to capture proteins with some level of binding affinity to the sequences on the column. The nucleic acid is often referred to as the bait, while the proteins that can bind to it are known as the prey. This is sometimes referred to as a pull-down method and is popular for the characterization of RNA or DNA binding proteins.

Other methods

The following procedure, a modification of Peppel and Baglioni (1990), Salvatori *et al.*, (1992), and Zolfaghari *et al.* (1993), exploits the ability of the ionic detergent SDS to inhibit RNase activity. It is a very rapid method for isolating total cellular RNA without guanidinium buffers, phenol, or isopycnic centrifugation, and was the basis for the development of a number of popular non-phenol-based kits for RNA isolation. It is suitable for the isolation of both eukaryotic and prokaryotic RNA. When RNA is to be isolated from tissue, it is best to couple this method with some form of mechanical disruption. Most of the reagents required are commonly found in labs performing even the most rudimentary molecular biology procedures. The volumes can be scaled up to accommodate larger extractions, and one may expect to harvest approximately 100 μg RNA per 10^6 cells.

Protocol: rapid isolation of RNA with SDS and potassium acetate reagents

1. Wear gloves!
- 2a. For tissue culture: Remove growth medium from the tissue culture vessel containing cells of interest and rinse with PBS. Remove PBS by aspiration and discard. Lyse cells directly on tissue culture plastic with lysis buffer (2% SDS; 200 mM Tris-Cl, pH 7.5, 0.5 mM EDTA), using 50 μl lysis buffer/cm². Gently agitate the culture vessel on the bench top or on an orbital platform for 2 min to ensure complete lysis. This lysis buffer can be prepared ahead of time and stored at room temperature.
- 2b. For tissue: *Rapidly* harvest the tissue and mince in ice-cold saline solution. Quickly transfer tissue into a suitable homogenization vessel that already contains an aliquot of lysis buffer (step 2a). If tissue was flash-frozen in liquid nitrogen, it may be helpful to shatter it (while wrapped in aluminum foil) by impact upon the lab bench. In so doing, the sample can be quickly transferred and allowed to thaw in the lysis buffer. Use 1 ml lysis buffer for each milligram of tissue. Disrupt the cells or tissue at room temperature by either manual homogenization (repeated shearing through a 19-gauge needle; Dounce homogenizer) or mechanical homogenization (e.g., Polytron) and avoid generating excess heat or foam during this step.
3. Transfer the lysate to a sterile microfuge tube (1.7- or 2.2-ml size). Cap and invert tube sharply 15–20 times.
4. For each 500 μl lysis buffer used in step 2, add 150 μl of potassium acetate solution (50 g potassium acetate; 11 ml glacial acetic acid; H₂O to 100 ml). This solution can be prepared ahead of time and stored at -20° .
5. Carefully invert the tube sharply 15–20 times to ensure thorough mixing.

6. Incubate on ice for at least 3 min but no longer than 5 min.
7. Centrifuge sample for 5 min at room temperature at maximum speed in a microcentrifuge.
8. Carefully recover the supernatant and transfer to a fresh tube.
9. Extract the supernatant with 300 μ l of a mixture of chloroform:isoamyl alcohol (24:1). Centrifuge for 30s to separate the phases. Transfer the aqueous (upper) phase to a fresh tube without disturbing the protein interface and then repeat the extraction.
10. Transfer the aqueous (upper) phase to a fresh microfuge tube pre-chilled on ice.
11. Precipitate the RNA with the addition of an equal volume (approximately 650 μ l) ice-cold isopropanol. Incubate at -20° for 30 min.
12. Collect the RNA precipitate by centrifugation for 5 min at room temperature.
13. Carefully decant the supernatant; wash the cell pellet 2–3 times with 70% ethanol. If the pellet becomes dislodged, centrifuge the sample again before attempting to remove the ethanol. Decant and discard the ethanol, and allow the sample to air-dry. If desired, a final wash with 95% ethanol may facilitate air-drying.
14. Dissolve RNA in the smallest possible volume of TE buffer or nuclease-free H_2O . Determine RNA concentration and store sample in suitable aliquots at -80° until further use.
15. Optional: Incubate purified sample with RNase-free DNase, as described in Appendix F, to remove any contaminating DNA.

Protocol: isolation of prokaryotic RNA

This rapid and efficient protocol for the isolation of bacterial RNA is a modification of the procedures of Peppel and Baglioni (1990), Salvatori *et al.* (1992), and Zolfaghari *et al.* (1993). Be sure to use the lysis buffer described below; *never* use the standard NaOH/SDS lysis buffer used to disrupt bacterial cells for plasmid preps, as this buffer will cause rapid and extensive hydrolysis of the RNA!

1. Wear gloves!
2. In advance: Prepare the NP-40 lysis buffer and potassium acetate buffer, and store them at 4° until ready to use. Although these buffers have a shelf life of several months, they should be inspected prior to use for the growth of opportunistic prokaryotes.

NP-40 Buffer

250 mM sucrose
20 mM EDTA, pH 8.0
0.75% NP-40⁹

Important: do not add the NP-40 detergent until *after* the other two components have been mixed and autoclaved.

Potassium Acetate Buffer (in order)

Prepare 3 M potassium acetate (KCH_3O_2) in 80% of final volume.
Adjust the pH to 5.5 with glacial acetic acid.

⁹Use IGEPAL CA-630 if NP-40 is not available.

Add remaining H₂O.

Add 100 g guanidinium thiocyanate (GTC) per 100 ml.

Mix carefully and thoroughly.

3. Harvest bacterial cells from a 100 ml culture by centrifugation. Completely remove supernatant.
4. Combine 1.5 ml of ice-cold lysis buffer with 2 ml ice-cold potassium acetate buffer, above, and then gently resuspend the cell pellet in this mixture. Allow lysate to incubate on ice for 10 min.
5. Centrifuge at $10,000 \times g$ for 10 min. Be sure to use centrifuge tubes rated for this *g*-force. It may be helpful to divide the lysate into several microcentrifuge tubes so that the investigator always has access to several aliquots of RNA made from the same culture on the same day.
6. Transfer the supernatant to a fresh tube, taking care not to transfer any of the precipitated solids. It is far better to leave a small amount of supernatant behind than to risk carryover of cell debris. Incubate supernatant on ice for an additional 15 min.
7. Centrifuge at $10,000 \times g$ for 10 min, at 4° if possible.
8. Transfer the supernatant to a fresh polypropylene tube, and then extract with an equal volume of chloroform:isoamyl alcohol (24:1). Mix *carefully* and thoroughly. Centrifuge for 3 min to separate the phases. If a protein band appears between the upper and lower phases, repeat the extraction with a fresh aliquot of chloroform:isoamyl alcohol (24:1).
9. Transfer the upper aqueous phase to a fresh tube, and then add 0.6 volume of ice-cold isopropanol. Store sample at -20° for 20 min to precipitate the RNA.
10. Collect RNA by centrifugation at $10,000 \times g$ for 10 min, at 4° if possible.
11. *Carefully* decant supernatant and then wash pellet with several aliquots of 70% ethanol. A final wash with 95–100% ethanol will accelerate the drying process.
12. Allow the sample to air-dry, and then resuspend in TE buffer or nuclease-free H₂O. After determination of RNA concentration and purity, store in suitable aliquots at -80°. If RNA is not to be used right away, store as an ethanol precipitate at -80°.

Protocol: isolation of RNA from yeast

Yeast produce a cell wall that acts as a formidable impediment to rapid recovery of RNA from these eukaryotic cells. Many of the classical methods for the isolation of RNA from yeast involve vortexing yeast cultures in the presence of glass beads, a technique intended to compromise severely the ultrastructure of these cells in a non-enzymatic manner.¹⁰ This type of approach is tedious and, when scaled down to the mini-prep level, often results in poor yields.

The RNA isolation procedure given here exploits the chaotropic nature of phenol and SDS, enhanced by heating, freezing, and then thawing the sample. As the sample freezes, the phenol crystals form which pierce the yeast cell,

¹⁰The purification of yeast genomic DNA often involves incubation of yeast cells with the enzyme zymolyase, the result of which is degradation of the cell wall and concomitant spheroplast formation. Such an approach results in gentle cellular disruption and favors the recovery of high-molecular-weight DNA, though it is not necessary for RNA isolation.

thereby liberating its contents. DNA is separated from RNA by phenol extraction under acidic conditions, a tactic frequently exploited for RNA purification under chaotropic conditions. The procedure presented here is a modification of the method of Schmitt *et al.* (1990). It has the advantages of being a rapid isolation procedure and having minimum cell culturing requirements.

1. *In advance:* Inoculate 0.5 ml of yeast culture medium (YPD)¹¹ in a sterile 15 ml tube. Grow culture overnight at 30° with gentle shaking.

Note: Ensure that the tubes are properly covered, not only to prevent evaporation but also to prevent infection by opportunistic prokaryotes.

2. The following morning dilute the culture 1:40 with fresh YPD medium. Continue to incubate for 4 hours or until the culture is in exponential growth phase.
3. Remove a 5 ml aliquot of the culture, and collect cells by centrifugation (500 × g for 5 min). Remove the supernatant as completely as possible and discard.
4. Resuspend cell pellet in 400 µl of AE buffer (50 mM sodium acetate, pH 5.2; 10 mM EDTA). If the cells do not resuspend quickly and thoroughly, then add more AE buffer. *Note: If desired, an appropriate RNase inhibitor may be added to the solution just prior to use, though this is usually not necessary.*
5. Transfer the cell suspension to a microfuge tube and add 40 µl of 10% SDS. Invert the tube several times to mix. *If the volume of AE buffer was increased in step 4, be sure to increase the volumes of all other reagents proportionally.*
6. Add an equal volume (approx. 1 ml) of molecular biology-grade phenol that has been equilibrated with AE buffer.

Important: Do not adjust the pH of the equilibrated phenol. See Appendix D for tips on phenol preparation.

7. Mix the sample carefully and thoroughly and then incubate at 65° for 5 min.
8. Rapidly cool the lysate on dry ice until phenol crystals appear. If an ethanol/dry ice bath is not available, the sample may be incubated at -20° for 30 min.
9. Centrifuge the sample at 12,000 × g (top speed) in a microfuge for 2 min to separate the phases.

Note: It is very common to observe a large amount of precipitate, consisting of cellular debris, at the bottom of the tube in addition to a protein interface between the upper aqueous phase and the lower organic phase.

10. Carefully recover the upper (aqueous) phase, taking care not to disturb either the protein interface or the organic phase. Transfer the aqueous material to a fresh microfuge tube and then add an equal volume of room temperature modified phenol:chloroform:isoamyl alcohol (25:24:1), prepared as described in Figure 2.2. Mix carefully and thoroughly by inversion for 2–3 min.
11. Centrifuge the sample at 12,000 × g in a microfuge for 2 min to separate the phases.
12. Carefully recover the upper (aqueous) phase, taking care not to disrupt either the protein interphase (if present) or the organic phase. Transfer the aqueous material to a new microfuge tube.
13. Extract the aqueous material once with an equal volume of chloroform or chloroform:isoamyl alcohol (24:1). Pulse centrifuge for 30 s to separate phases and then carefully transfer the aqueous phase to a fresh tube.

¹¹YPD medium (per liter) = Mix 10g yeast extract, 20g peptone, 20g dextrose; autoclave. After autoclaving, YPD medium acquires a dark brown appearance due to the “caramelizing” of the dextrose at high temperatures.

14. Add 0.1 volume 3M sodium acetate, pH 5.2, and 2.5 volumes of 95% ethanol. Store sample overnight (or at least for a few hours) at -20° to precipitate RNA.
15. Collect the precipitate by centrifugation at $12,000 \times g$ for 8 min, and do so at 4° if possible.
16. Decant ethanol. Wash the pellet 3 times with $500\mu\text{l}$ 70% ethanol. Air-dry the pellet, but *do not* allow the pellet to dry completely. If desired, a final wash with 95% ethanol may accelerate air-drying the RNA pellet. Store the sample at -80° until ready to use.
17. Resuspend the pellet in $25\text{--}50\mu\text{l}$ nuclease-free H_2O or TE (10mM Tris, 0.1 mM EDTA; pH 7.5).
18. Determine RNA concentration and purity as described in Chapter 6. Store the sample in aliquots at -80° . The expected yield is about $150\mu\text{g}$ total RNA from an exponential 10 ml yeast culture.
19. As always, be sure to electrophorese an aliquot of the newly isolated RNA to ascertain the integrity of the sample (Chapter 9).

Short- and long-term storage of purified RNA

The correct storage conditions for RNA samples is often a hotly discussed topic within the laboratory. Improper storage, over a period of merely a few hours or as long as several months, is likely to have a profound negative impact on the probable utility of a sample of purified RNA. The key issues are:

- a. Has the RNA been purified from the cellular or tissue architecture?
- b. Is -80° storage available in the lab?
- c. Will the RNA be used within the next 7–10 days?
- d. Has the RNA been redissolved in aqueous buffer (nuclease-free water or TE buffer)?

Tissues and cell pellets may be harvested and then flash-frozen for storage in liquid nitrogen or at -80° . When stored as such, RNA can be purified up to a year later. In general, the frozen material is thawed in lysis buffer, accompanied by immediate mechanical disruption. Other investigators will homogenize fresh tissue in guanidinium buffer upon receipt and then freeze the homogenate at -80° , continuing the RNA isolation procedure at a later date.

Purified RNA is most stable when stored as an ethanol precipitate at -80° . Under these conditions RNA can be stored confidently for several months or even longer, the half-life being a direct function of the biological source.¹² If the sample will surely be used within the next week, the RNA may be stored stably at -20° , again as an ethanol precipitate. The “clock begins ticking” when a purified sample of RNA is dissolved in aqueous buffer, either sterile water or modified TE buffer (10 mM Tris; 0.1 mM EDTA; overall pH 7.5). TE buffer has the advantage of being able to chelate the Mg^{2+} that is often carried

¹²RNA purified from certain sources (plants, for example) is remarkably stable for years, when maintained as an ethanol precipitate at -80° . These RNA species also show enhanced stability when stored for weeks in hydrated form at -20° .

over from certain isolation procedures, while the low EDTA concentration is unlikely to cause havoc in downstream PCR procedures (discussed in Chapter 18). Magnesium and other heavy metals are known for their ability to catalyze non-specific degradation of RNA, so chelation is advantageous.

RNA stored in nuclease-free H₂O or TE buffer should be maintained on ice during handling, unless a protocol specifically dictates otherwise. The most prudent course of action would be to determine the RNA concentration and then store the RNA in suitable aliquots at -80° . This will preclude repeated freezing and thawing of any one tube. Some investigators add RNase inhibitors to RNA samples during long-term storage, an action that is usually unnecessary and may even be counterproductive. It is incumbent upon the investigator to ensure that any RNase inhibitors added to the prep at any time will not interfere with any subsequent manipulations and/or reactions involving the RNA.¹³

Other options for long term storage of RNA are (1) highly purified 100% formamide (Chomczynski, 1992); (2) a commercially available stabilized form of formamide known as FORMAzol[®] (Molecular Research Center, Cincinnati, OH), or (3) RNAlater (Ambion, Austin, TX), which is a proprietary formulation for stabilizing RNA within tissue samples and cell pellets prior to isolation.

RNA is stable for up to two years at -20° when stored in FORMAzol, and this reagent has been used very successfully in this lab. The RNA can then be used directly for denaturing gel electrophoresis (Chapter 9). If required for cDNA synthesis or RT-PCR, either formamide or FORMAzol are easily removed by precipitating the RNA with the addition of 4 volumes of 95% ethanol; at dilute concentrations, the addition of 0.2 M NaCl may also be required for efficient recovery of the RNA.

There are many considerations with respect to the correct storage of RNA. If ever in doubt, store purified RNA samples, in whatever form, at -80° .

References

- Carter, C., V.G. Britton and L. Haff. (1983). CsTFA: a centrifugation medium for nucleic acid isolation. *BioTechniques* 1, 142–147.
- Castora, F.J. and J.J. Greene (Eds). (2009). *Handbook of Biological Research*. CRC Press, Boca Raton, FL.
- Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald and W.J. Rutter. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18, 5294–5299.
- Chomczynski, P. (1992). Solubilization in formamide protects RNA from degradation. *Nucleic Acids Res.* 20, 3791–3792.
- Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA, and proteins from cell and tissue samples. *BioTechniques* 15, 532–536.

¹³Historically, RNA has been stored in widely varying concentrations of SDS, VDR, and ultrapure formamide. Unfortunately, the incomplete removal of these “protectant” compounds prior to use will severely reduce the utility of the RNA in most downstream applications.

- Chomczynski, P. and N. Sacchi. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.* **162**, 156-159.
- Cooper, T.G. (1977). *The Tools of Biochemistry*. John Wiley & Sons, Inc, New York (pp. 309-354).
- Cox, R.A. (1968). The use of guanidinium chloride in the isolation of nucleic acids. In "Methods in Enzymology," (L.Grossman & K. Moldave, Eds.), Vol 12, part B, pp.120-129. Academic Press, Orlando, FL.
- Favaloro, J., R. Triesman and R. Kamen. (1980). Transcriptional maps of polyoma virus-specific RNA: analysis by two-dimensional S1 gel mapping. *Methods Enzymol.* **65**, 718-749.
- Glšičin, V., R. Crkvenjakov and C. Byus. (1974). Ribonucleic acid purified by cesium chloride centrifugation. *Biochem.* **13**, 2633-2637.
- Gordon, J.A. (1972). Denaturation of globular proteins. Interaction of guanidinium salts with three proteins. *Biochem.* **11**, 1862-1970.
- Lindberg, U. and T. Persson. (1974). Messenger RNA isolation with poly(U) agarose. *Methods Enzymol.* **34**, 496-499.
- Majumdar, D., Y.J. Avissar and J.H. Wyche. (1991). Simultaneous and rapid isolation of bacterial and eukaryotic DNA and RNA: A new approach for isolating DNA. *BioTechniques* **11**(1), 94-101.
- Nozaki, Y. and C. Tanford. (1970). The solubility of amino acids, diglycine, and triglycine in aqueous guanidinium hydrochloride solutions. *J. Biol. Chem.* **245**, 1648-1652.
- Okayama, H., M. Kawaichi, M. Brownstein, F. Lee, T. Yokata and K. Arai. (1987). High efficiency cloning of full length cDNA: Construction and screening of cDNA expression libraries from mammalian cells. *Methods Enzymol.* **154**, 3-28.
- Peppel, K. and C. Baglioni. (1990). A simple and fast method to extract RNA from tissue and other cells. *BioTechniques* **9**, 711-713.
- Salvatori, R., R.S. Bockman and P.T. Guidon, Jr. (1992). A simple modification of the Peppel/Baglioni method for RNA isolation. *BioTechniques* **13**, 510-511.
- Schmitt, M.E., T.A. Brown and B.L. Trumpower. (1990). A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **18**(10), 3091-3092.
- Ullrich, A., J. Shine, J. Chirgwin, R. Pictet, E. Tisher, W.J. Rutter and H.M. Goodman. (1977). Rat insulin genes; construction of plasmids containing the coding sequences. *Science* **196**, 1313-1319.
- Zarlenga, D.S. and H.R. Gamble. (1987). Simultaneous isolation of preparative amounts of RNA and DNA from *Trichinella spiralis* by cesium trifluoroacetate isopycnic centrifugation. *Analyt. Biochem.* **162**, 569-574.
- Zolfaghari, R., X. Chen and E.A. Fisher. (1993). Simple method for extracting RNA from cultured cells and tissues with guanidine salts. *Clin. Chem.* **39**, 1408-1411.

3 The Truth about Tissues

Rationale	81	Protocol: LiCl-Urea method for RNA isolation from tissue	92
Tissue culture or tissue?	81	Protocol: RNA isolation from lipid-enriched samples	95
Advantages of cell culture	82	Purification of polysome-engaged mRNA	96
Advantages of tissue samples	83	Protocol: isolation of polysomal mRNA	98
Homogenization methods	84	Collecting samples in the field	99
Polytron disruption	84	RNA "Clean-Up" methods	100
Dounce homogenization	86	Troubleshooting RNA isolation from tissue	101
BeadBeater™ technique	87	References	102
RNA isolation strategies for various organs and tissues	87		
Fresh tissue	89		
Frozen tissue	90		
Fixed tissue	91		

Rationale

There is much to be said in favor of a good *in vitro* model for both basic- and applied research. However, it is nearly impossible to extrapolate results obtained *in vitro* to completely and accurately describe *in vivo* phenomena. Unquestionably, the ability to detect the many subtle and not-so-subtle biochemical changes that accompany a natural phenomenon or experimental manipulation is lost when working with cell culture. In contrast, working with cells grown in culture presents fewer logistical difficulties than the acquisition of tissue samples or working with whole organism models, both animals and plants. Succinctly, there is much to be said in favor of developing a cell culture model; likewise there is much to be said against it. This chapter addresses issues unique to animal tissue while the isolation of RNA from plants tissue is the focus of Chapter 4.

Tissue culture or tissue?

The isolation of high quality RNA from cells grown in culture is generally far easier to accomplish than is an attempted RNA isolation from tissue. The primary

reason is that the three-dimensional architecture of the tissue that must be dismantled and removed compared to disruption of a cell monolayer¹. Frequently, aggressive methods of homogenization in the presence of rather harsh formulations of lysis buffers are required to provide a satisfactory level of tissue disruption so as to render an RNA yield of acceptable quantity and purity. Further, some tissues are notorious for their legendary levels of RNase and reluctance to give up their RNA. Thus, the aggressive methods needed to compensate for the unique character of tissue often results in the isolation of total cellular RNA, as opposed to a subpopulation of RNA, i.e. cytoplasmic *or* nuclear transcripts, whether intended or not. The good news is that tissues can be snap-frozen, or otherwise stabilized, immediately after harvesting, and stored at -80° for isolation at a more convenient time.

RNA isolation from tissue is a non-negotiable when studying diseased tissue, when attempting to ascertain a tissue-specific response to an experimental challenge or other insult, or when trying to characterize the molecular physiology of normal tissue. In addition to some of the RNA isolation difficulties described previously, the tissue sample itself may be difficult to acquire especially when the “phenomenon” is rare or when patient consent is required. Investigators working with human samples also know all too well of the inherent dangers of HIV- or hepatitis-tainted samples, to name but two prevalent, extremely dangerous infectious agents. Thus, not only are the mechanics of RNA isolation from tissue samples modified, compared to RNA isolation from cell culture samples, so is the level of difficulty in the storing and handling of the sample so that the safety of all parties involved is maintained.

The following list summarizes some of the advantages and disadvantages associated with the use of cell culture and tissue samples, respectively.

Advantages of cell culture

1. The total exposed surface area of cells grown in culture is much greater than is observed in a tissue sample containing an equivalent number of cells. Consequently, it is easier to ensure complete cell lysis. In contrast, some tissues are extremely difficult to disrupt, while other tissues are quite recalcitrant to giving up their RNA.
2. Many primary cells and established cell lines are easy to culture and can be propagated easily and economically.
3. Cells growing in culture are available on an as-needed basis. In contrast, it is often difficult to obtain tissue samples or biopsies from animals on a regular basis and, especially with humans, nearly impossible to acquire multiple tissue specimens from the same organism.
4. Isolation of RNA from cells in culture is generally more expedient than isolating the same from a tissue sample.

¹While many transformed cell lines have lost the capacity for contact inhibition and thus form readily observable foci *in vitro*, the investigator usually is dealing with a relatively small number of cells piled up on top of each other, rather than the millions of cells held together by formidable intercellular junctions and the extracellular matrix (ECM) *in vivo*.

5. Disruption of tissue samples generally requires fairly aggressive homogenization, including Polytron® disruption or Dounce homogenization, an approach not associated with the disruption of cells grown in culture. If excessive heat is generated, then the stability of the RNA will be compromised.
6. The mechanical shearing forces that accompany tissue disruption usually disrupt membrane-bound organelles, including the nucleus. Thus, tissue homogenization should be expected to yield routinely significant quantities of nuclear RNA, genomic DNA, and mitochondrial and/or chloroplast DNA. This will necessitate additional steps to clean up the RNA prior to use. In contrast, it is very easy to isolate RNA from specific subcellular compartments when working with cell culture.
7. Cells grown *in vitro*, once removed from the culture vessel, generally dissociate from each other with minimal difficulty. However, cells that constitute a tissue are tethered together via various types of intercellular junctions as well as components of the extracellular matrix, or ECM. Components of the ECM contribute to the heavy debris that commonly accompanies tissue disruption, thereby necessitating one or more additional steps to remove completely this debris in order to effect recovery of RNA that is clean enough to be useful. While these additional required steps are not technically challenging, they do constitute additional manipulations and, concomitantly, more opportunities to lose or compromise the sample.
8. Cells grown in culture usually represent a homogeneous population. In contrast, even a very small tissue snippet may contain many types of cells, making it difficult to pinpoint the specific cells responsible for any observed modulation of gene expression.

Advantages of tissue samples

1. The three-dimensional architecture of the tissue is preserved. This permits thin sectioning of flash frozen tissues with a cryostat or with a microtome in the case of formalin-fixed samples for localization of gene expression by *in situ* hybridization. Alternatively, a newer technology known as laser capture (Fend *et al.*, 1999; Kuecker *et al.*, 1999; Suarez-Quian *et al.*, 1999) can be applied to tissue samples for the isolation of RNA from a very small number of specific cells within the tissue.
2. The natural molecular physiology of the tissue is preserved, whether normal, wild type, untreated, mutant, or treated.
3. Relatively small tissue samples (100 mg range) can yield sufficient RNA for multiple applications.
4. Depending on the mass of a tissue sample, a very large quantity of RNA can be isolated because of the enormous numbers of cells involved. Storing RNA from a large-scale isolation ensures the availability of identical aliquots of RNA from multiple experimental replicates.
5. The stability of RNA in a tissue sample is excellent when the tissue is flash-frozen in liquid nitrogen and then stored at -80° , even for extended periods. Ultracold tissue samples can be fractured into smaller pieces or pulverized for later use. When ready to perform RNA isolation, thawing the tissue directly in the presence of the lysis buffer has always worked very well in this lab.
6. The results of transcription-based assays when starting with tissue samples reflect the pharmacology of a drug on a specific organ or system.

Homogenization methods

One of the most common mistakes that investigators make when working with tissue for the first time is the use of too much tissue at the onset of the RNA isolation procedure. This is especially true if cell culture was used previously in the lab as a biological source for RNA isolation. If the lysis buffer is overloaded with too much tissue, then the purity of the RNA is likely to suffer and the overall yield will plummet. If, on the other hand, the volume of lysis buffer is increased, it may be necessary to split the homogenate into multiple tubes, making processing of the sample cumbersome at best, and reducing the number of different samples that can be handled simultaneously.

Two basic approaches to homogenization are presented here, the choice of which depends upon the difficulty encountered in disrupting the tissue type under investigation. Keep in mind that animal tissue is generally easy to disrupt because the cells lack the cell wall associated with plants, bacteria, and yeast. As a general rule, one should use the least aggressive means necessary for creating the homogenate. The idea behind homogenization is simply to increase the surface area of the biological material to afford the lysis buffer a chance to do its job: lyse cells. Overly zealous investigators endeavoring to disrupt every single organelle by mechanical means run the risk of damaging the RNA to the point where it will no longer support downstream analyses.

Polytron disruption

The term *polytron* enjoys commonplace usage in most research laboratories. Polytron[®] is a registered trademark-protected name (Kinematica, Inc., Bohemia, NY) of a motorized device used to rapidly homogenize all types of plant and animal tissue, as well as bacterial cultures, in large and small volumes (Fig. 3.1). Attached to the motor is a stainless steel probe, properly known as a “generator”. The end of the generator is characterized by very sharp, knifelike, stainless steel saw-tooth blades that are designed to shred tissue efficiently. The variable-speed motor allows that user to control the degree of tissue disruption while minimizing the amount of foaming and heat generation. It is extremely important for the user to observe all safety procedures recommended by the manufacturer, as intentional misuse or accidents may result in serious, permanent injury. Motorized homogenization devices are also available from other companies such as Omni International (Marietta, GA). To a great extent, the use of hand-held and stationary homogenizers for nucleic acid work has replaced the use of the Waring blender which, at one time, was a laboratory staple.

Although stainless steel generators are durable and may be required to homogenize recalcitrant tissues, a newer generation of disposable generators was introduced in the late 1990s by Omni International (Fig. 3.2). The use of disposable generators dispenses with the need to clean a single stainless steel probe when multiple samples must be processed. This approach eliminates the



Figure 3.1 Polytron[®] homogenizer. Models range from hand-held units (shown) to dispersing machines with throughput capacity up to 130,000 L/h. Courtesy of Kinematica, Inc.



Figure 3.2 Hand-held motorized homogenizer showing disposable generator tips. Tips are simply discarded after use, minimizing the down-time required to clean generator between each sample. If desired, the tips can be cleaned and reused a limited number of times. Courtesy of Omni International.

potential for cross-sample contamination and is particularly helpful in the control of RNase activity, thereby favoring maximum stability of the RNA. These disposable homogenizer generators have been used extensively and with great success in this lab for soft-tissue homogenization. In the case of human tissues, the use of disposable generators also reduces the potential for the spread of

infectious agents. The company reports that their plastic generator probes can be autoclaved as many as seven times if re-use is desired.

Dounce homogenization

The Dounce homogenizer is a hand-held glass implement that resembles a test tube (Fig. 3.3). Generally the open end of the instrument is enlarged to some degree to accommodate the introduction of biological material. Inserted into the glass tube is a stylized tightly-fitting round glass pestle with a grip at one end. This design allows the user to apply a desired number of up-and-down strokes while twisting the pestle in the process; this motion maximizes the shearing forces. Dounce homogenization is considered a form of liquid homogenization because tissue disruption occurs as the sample is forced through a narrow space, specifically between the inside wall of the homogenizer and the pestle. In contrast to some of the physical (mechanical) forms of cell and/or tissue disruption that often require the use of expensive equipment, Dounce homogenization is a much more gentle and economical means of breaking up tissue. A Dounce homogenizer is ideal for the isolation of intact nuclei and other organelles that would otherwise be compromised by aggressive mechanical disruption.

The traditional Dounce homogenizer is designed to be used over and over again. In this laboratory, glass homogenizers are thoroughly scrubbed and rinsed, and then soaked in 3% H_2O_2 (see Chapter 7) for 20 minutes before



Figure 3.3 Dounce homogenizer. This handheld instrument is an effective homogenization device for many types of tissue. See text for details. Courtesy of Bellco Glass Company.

being rinsed with copious amounts of autoclaved water. After air-drying, each homogenizer and its pestle are individually wrapped in foil and autoclaved to ensure that the homogenizer is nuclease-free the next time it is used.

One may also purchase disposable single-use conical tissue grinders. These devices are Dounce-like homogenizers that consist of a plastic pestle mounted in the cap of a conical 50 ml or 15 ml tube; small versions are also available for use with 1.5 ml microfuge tubes. These devices are extremely handy when large numbers of samples need to be processed in a short span or to control the dispersal of infectious agents that could potentially be in the sample.

BeadBeater™ technique

The BeadBeater product line (BioSpec Products, Inc., Bartlesville, OK) is innovative instrumentation in widespread use for the disruption of cells that have a cell wall. This is an economical method that can be used when chemical lysis is inefficient or when chemical lysis yields are unacceptably low. The BeadBeater (Fig. 3.4) approach is a mechanical lysis method in which small beads (smooth or sharp) are used to disrupt the cells rapidly and in various types of lysis buffer. Bead size is selected based on the material to be disrupted, and typically ranges from a diameter of 0.1 mm to more than 2.0 mm. Several bead materials and densities are also available, depending on the nature of the biological material.

The method of cell lysis is based on the crushing action that occurs when the beads collide with the sample. By limiting the mechanical disruption to short bursts of the instrument and when used in conjunction with the accompanying cooling jacket, the BeadBeater method facilitates high-yield macromolecule recovery. Through the use of relatively gentle lysis buffers, for example, large quantities of functional proteins, enzymes, and nucleic acids may be recovered. The method is also compatible with the isolation of small organelles. While the beads used for disruption may be cleaned and used again, most investigators simply discard them to preclude any possible cross-contamination among samples and because the beads are quite inexpensive.

The BeadBeater approach is also used extensively for the isolation of nucleic acids from microorganisms found in environmental samples, including soil and water. The yield of both DNA and RNA from samples processed in this manner is generally much greater than through the use of chemical extraction alone. Other instrumentation for mechanical disruption of this nature are the FastPrep® system (MP Biomedicals; Solon, OH) and the Precellys® homogenizer (MO BIO Laboratories; Carlsbad, CA).

RNA isolation strategies for various organs and tissues

There is no single fool-proof method for the expedient isolation of RNA from tissue samples. Each tissue type has qualities that make the isolation methodology

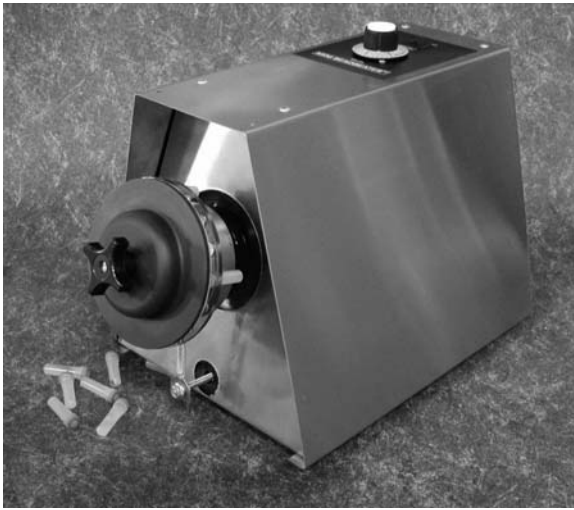


Figure 3.4 Mini-BeadBeater 16. This is a high-energy cell disruptor that can process up to 16 samples simultaneously. Courtesy of BioSpec Products, Inc.

unique to that sample. The tissue mass, the cell type(s) involved, the age of the tissue, the status of the tissue (fresh, flash-frozen, formalin-fixed, animal or plant) are all important considerations. No matter what the source, the key concern is the stability of the RNA in the tissue until the RNA isolation procedure commences.

One guiding principle when planning to isolate RNA from tissue samples: it is *not* necessary to accomplish complete tissue disruption in order to recover acceptable quantities of usable RNA. Being overly enthusiastic, i.e. trying to break up every single clump of tissue, may be to the detriment of the resulting RNA in terms of both quality and quantity. If the tissue breaks up quickly and efficiently, fine. If the tissue does not dissociate efficiently, remember that even moderate disruption results in an increased surface area on which the lysis buffer can act. This may be all that is necessary to isolate enough RNA for blot analysis, RT-PCR, or related applications. Further, leaving a partial homogenate in the presence of the lysis reagent for too long at room temperature can be completely counterproductive, meaning that the RNA quality declines with increased exposure to the lysis conditions². In other words, when working with tissues, timing is everything.

²In some laboratories tissue samples are homogenized on receipt in one of the guanidinium based buffers, followed by storing the homogenate at -80° without any further purification. This approach is generally acceptable for most tissues, though some labs have experienced difficulty recovering high quality RNA after the homogenate was thawed and the RNA isolation procedure completed. There really are no hard-and-fast rules, as the success of this strategy is both tissue- and user-dependent.

One common mistake made when working with tissue samples for the first time involves that actual mass of tissue that is used. In general, most of the protocols that appear in the literature or accompany commercial reagents for RNA isolation are optimized for 100 to 1000 mg of tissue. When too much tissue is used without scaling up reagent volumes appropriately the RNA yield almost always declines. Unfortunately, scaling-up often results in a final volume that is too large for a single tube, resulting in the extreme inconvenience of having to perform multiple centrifugation steps. At the same time, the tissue should be in small enough pieces so as to be able to interact with the lysis buffer and the homogenizer itself. Large chunks of tissue, including soft tissue, do not homogenize well and often clog the homogenizer, especially the motor-driven hand-held homogenizers that are in wide-spread use.

One of the unique features of successful RNA isolation from tissue samples is the fact that very large quantities of RNA are often recovered at the end of this procedure, depending on the input mass and the type of tissue. This may require as much as 500 to 1000 μ l of sterile buffer to dissolve the RNA. For investigators accustomed to working with RNA from cell culture, this required large volume may be quite unsettling. Remember that even small tissue samples consist of a surprisingly large number of cells, yielding a proportional mass of RNA. If the RNA is not dissolved completely, it will appear electrophoretically as a massive smear (Fig. 3.5), indicating that more buffer is required. In such an instance, it is wise to take small, representative aliquot from the master tube, and then dilute only the aliquot until the 28S and 18S rRNAs become plainly visible upon denaturing electrophoresis, as described in Chapter 9. After the concentration of the now-diluted aliquot has been determined spectrophotometrically, the remaining RNA can be diluted down to a pre-determined standard storage concentration and then, if desired, reprecipitated in convenient aliquots for long-term storage at -80° . At some point prior to the completion of this procedure, it is very important to perform some of the quality control methods described in Chapter 6 because of the added complexity of many RNA isolation procedures involving tissue, compared to cell culture.

Fresh tissue

It is imperative to realize that RNase activity remains a viable weapon of mass destruction after a tissue sample has been removed from the donor organism. Further, RNase remains active following cell disruption as long as conditions persist that support the function of these recalcitrant enzymes. The idea is to disrupt the tissue as quickly as possible and, in so doing, inactivate pesky ribonuclease. Immediately upon harvesting, the tissue should be rinsed in ice-cold phosphate-buffered saline (PBS), weighed, and then rapidly minced into small pieces using a nuclease-free scalpel. It is wise to avoid reusable scalpel blades, especially if the history of the scalpel is unknown; a fresh disposable blade should be used for each tissue sample. The minced tissue is then dropped into lysis buffer as quickly as possible. When transferring minced tissue into lysis

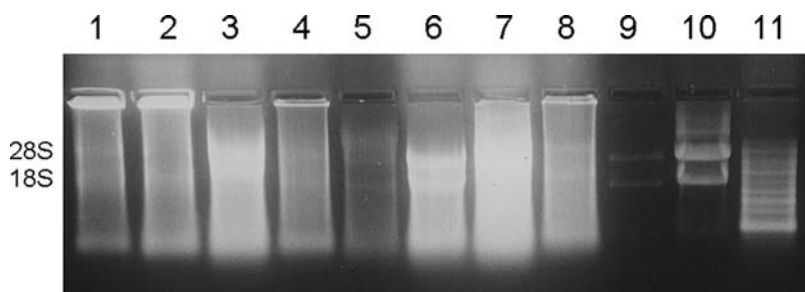


Figure 3.5 Appearance of incompletely dissolved RNA. Lanes 1–8: RNA was isolated from various human tissues and precipitated. The resulting clear, jelly-like pellet was diluted in 500 μ l TE buffer, incubated at 60° for 10 min, and then stored on ice for 2 h. RNA samples were electrophoresed through a denaturing formaldehyde-agarose gel, as described in Chapter 9. Since the amount of RNA recovered from the tissue was so great, 500 μ l was not enough to completely dissolve it, which also compromised the accuracy of spectrophotometric concentration measurements. Aliquots of the RNA samples were removed from the master tube and diluted further until the 28S and 18S rRNAs were plainly visible upon electrophoresis. Lanes 9–10: RNA that was isolated from human tissue at an earlier date and dissolved properly. Lane 11: RNA molecular weight standard (Promega). The samples shown in lanes 1–8 eventually demonstrated properly formed 28S and 18S rRNA bands after appropriate dilution.

buffer, it is best to transfer as little PBS as possible; if the lysis buffer is diluted too much it will lose its ability to disrupt the architecture of the tissue and, even worse, to inhibit RNase efficiently.

Frozen tissue

Tissue samples that are not processed immediately can also be frozen for analysis at a later date. Assuming that tissue samples are properly handled, meaning flash-frozen³ in liquid nitrogen as quickly as possible after ablation and then maintained at –80° or on dry ice until ready for use, there should be no difficulties isolating RNA of sufficient quantity and quality to support contemporary methods of analysis. Fresh, frozen tissue is then pulverized and added to chaotropic lysis buffer or, if the sample is small enough, allowed to simply thaw in a suitable aliquot of lysis buffer and then mechanically disrupted. However, if the tissue is to be sectioned on a cryostat for histological examination or *in situ* hybridization, most tissues section much better when they are embedded in O.C.T. compound⁴ and then frozen in the cryostat itself, rather

³The flash-freezing process is also known as “snap-freezing”.

⁴O.C.T. (optimal cutting temperature) is also known as Tissue-Tek® O.C.T. tissue embedding medium. This is a standard material used to prepare tissue samples for thin sectioning on a cryostat. O.C.T. is also used to attach the frozen tissue to the chuck, holding it in place while it is advanced toward the cryostat blade. Fresh tissue samples are generally cooled gradually in O.C.T. medium within the cryostat chamber at –22°.

than flash-frozen. Ensuring that the spatial distribution of the RNA is maintained is an absolute requirement if *in situ* hybridization experiments are to be meaningful; flash-freezing often distorts the spatial arrangement of certain tissue regions, thereby rendering unreliable data with respect to the local distribution of the transcripts of interest within specific cells or regions of the tissue.

Once tissue has been frozen, by any method, it is generally advisable to avoid thawing the tissue more than once. Freezing also makes subdividing or mincing the tissue extremely difficult without compromising the stability of the RNA. In this lab, small-to-medium size samples of flash frozen tissue (0.5–2 g) are twice wrapped in heavy-duty foil and stored in Ziploc® plastic bags at -80° . When ready to perform the RNA isolation, the sample is removed from ultracold storage and immediately whacked onto the surface of the lab bench while still frozen and wrapped⁵. The resulting impact generally shatters the tissue quite satisfactorily. Upon rapid unwrapping, shards of the frozen tissue are quickly transferred to a weigh boat that was pre-cool by placing it on dry ice. Thus, the mass of the biological material is measured and recorded prior to thawing the sample which, ideally, should be done in lysis buffer. It is usually a big mistake to allow frozen tissue to thaw prior to submersion in lysis buffer. Because the lysis buffer itself is not frozen, the immersed tissue will thaw very quickly in a strongly denaturing environment and mechanical disruption can commence immediately.

Fixed tissue

A well-established method for preserving tissue samples is fixation in 10% buffered formalin or 4% paraformaldehyde. While formalin fixation is commonplace, it has the undesirable consequence of making RNA recovery and analysis extremely difficult and inefficient. One reason for this difficulty has been the limited solubility of formalin-fixed tissues in guanidinium-based reagents. However, incubation with proteinase K is an effective method for solubilizing formalin-fixed tissues (Stanta and Schneider, 1991; Masuda *et al.*, 1999), thereby facilitating RNA and DNA recovery (Coombs *et al.*, 1999).

The RNA is able to support amplification PCR after removal of the monomethylol groups (-CH(2)OH) that are added to all four bases during the formalin-fixation process. These groups can be removed by incubating formalin-modified RNA in TE buffer (Masuda *et al.*, 1999) or, more recently, using a three-day lysis method at 65° (Chung *et al.*, 2006); in the case of the latter, amplicons in the range of 300 bp are routinely generated because the purity of the RNA recovered from formalin-fixed paraffin-embedded (FFPE) samples is good enough to support efficient reverse transcription. cDNA synthesized from RNA that was previous locked up in fixed tissue is usually smaller, compared

⁵Be sure to wear safety glasses, gloves, and be gowned according to departmental policy before performing this technique, and do be sure to check first to determine if it is allowable at all. Moreover, human specimens that could contain infectious agents should be handled in a suitable containment area. This caveat applies to other tissue as well.

to the situation where the RNA template is isolated from fresh- or flash-frozen tissue, though modified approaches to cDNA synthesis and PCR primer design have been demonstrated to improve the sensitivity of RT-PCR using formalin-fixed RNA as template (Mikhitarian *et al.*, 2004). As a general rule, it is much easier to extract and work with RNA from fresh or frozen tissue compared to tissue that has been fixed.

The ability to recover RNA from archival tissue offers an unparalleled opportunity to study unique samples that have been stored for perhaps as long as several years. This prospect is particularly tantalizing because of the substantial archives found in most clinical settings. This ability may provide new insight into disease progression, not to mention the ability to examine the genomes of both DNA and RNA viruses that may have been trapped in a tissue sample the moment that it became formalin-fixed. Because some viruses are known to mutate at an astonishing rate, one might be able to generate a genetic time line by which to compare contemporary viral strains with their predecessors.

Protocol: LiCl-Urea method for RNA isolation from tissue

The following is a modification of the procedures of Auffray and Rougeon (1980) and Kato *et al.* (1987). It has been used successfully for the isolation of RNA from a variety of tissues including breast, pancreas, liver, prostate, spleen, and brain. The key to success in this procedure is to homogenize the tissue in short bursts totaling no more than 60 s of homogenization time and do so while submerging the tube containing the tissue and lysis buffer in an ice water bath (Fig. 3.6). An earlier version of this protocol included an incubation with proteinase K, which has been omitted from this protocol. In this laboratory, the LiCl-urea method has produced large quantities of very pure RNA which often exceed the quality of RNA produced using guanidinium-based methods.

In advance: Make sure that all buffers and tubes are prechilled on ice in order to minimize nuclease activity and stabilize the RNA.

1. Place a weigh boat on dry ice for a few minutes to cool it down. This will allow accurate weighing of tissue without the sample thawing.
2. Rapidly transfer a suitable mass of the tissue into a prechilled 50 ml tube and then add 10 vol of LiCl buffer (3 M LiCl; 6 M urea; 50 mM Tris, pH 7.4; 1 mM EDTA; 100 mM β -mercaptoethanol, 0.05% sarcosyl).

Note: Add the β -mercaptoethanol just prior to use.

3. Immerse the tube in an ice water bath and homogenize the sample with several short bursts with a Polytron® or other tissue homogenizer. The total homogenization time should be no more than 60 s.

Note 1: It is very important to minimize foaming and the generation of heat during the homogenization process.

Note 2: Be sure to observe all safety precautions throughout the procedure when using a motorized homogenizer.

4. Transfer the homogenate to a prechilled polypropylene conical 15 ml tube (no more than 12 ml per tube). Place the tubes containing the homogenate at -20° overnight.
5. The following morning, place the tubes into a precooled rotor and centrifuge for 75 min at $5000 \times g$ at 4° .



Figure 3.6 Homogenization of tissue using the LiCl-urea method. The conical 50 ml tube containing the tissue sample and the LiCl-urea homogenization buffer is held in an ice bath during the homogenization process. This is an excellent way to avoid generating heat caused by the action of the homogenizer. Short bursts of homogenization are applied, totaling no more than 60s, as described in the text. The tube is maintained in the ice bath between applications of the homogenizer.

Note: Be sure to check the maximum relative centrifugal force (RCF) and chemical compatibility of the tubes being used in this step.

6. At the conclusion of the run, carefully decant and discard the supernatant, and then rapidly blot each tube dry with a Kimwipe. Observe the tube for the presence of solidified lipid clinging to the inside wall and, if present, wipe it away with a sterile cotton swab. Do not use a Kimwipe to try to remove anything from the inside walls of the tube. All tubes should be maintained on ice during this manipulation.
7. Add 8–10 ml cold 3 M LiCl to each tube, and gently vortex for 30s.

Note: The vortexing step should be only aggressive enough to disperse the pelleted material.

8. Centrifuge the samples for 45 min at $5000 \times g$, at 4° .

Note: Be sure to check the maximum RCF and chemical compatibility of the tubes being used in this step.

9. Repeat steps 6 to 8.
10. Resuspend the pellet in 2.5 ml of TE/SDS buffer (10 mM Tris, pH 7.5; 0.1 mM EDTA, 0.5% SDS). Vortex as needed to resuspend the pellet. If necessary, use a sterile serological pipette to disperse the pelleted material.
11. Add 150 μ l 5 M NaCl and mix by gentle vortexing.
12. If the material from step 11 is in a 15 ml polypropylene tube, then add an equal volume of modified phenol:chloroform:isoamyl alcohol (see note below). If the sample from step 11 is not in a polypropylene tube, it must be transferred to a tube that is compatible with these organic solvents. Secure the cap and then carefully mix the sample by vortexing.

Note: Modified phenol:chloroform:isoamyl alcohol is prepared by saturation molecular biology grade phenol with an equal volume of Tris-SDS buffer (100 mM NaCl; 1 mM EDTA; 10 mM Tris, pH 8.5; 0.5% SDS) and then adding a volume of chloroform:isoamyl alcohol (24:1) equal to the volume of the phenol/Tris buffer mixture.

Finally, add 8-hydroxyquinoline to final concentration of 0.1%. See Appendix D for additional details pertaining to phenol saturation.

13. Centrifuge samples at $1000 \times g$ for 10 min at 4° .
14. Transfer the upper aqueous phase to a fresh polypropylene tube. Be careful to avoid the protein interface.
15. Add an equal volume of modified phenol:chloroform:isoamyl alcohol. Carefully mix the sample by vortexing.
16. Centrifuge the sample at $1000 \times g$ for 10 min at 4° .
17. Transfer the upper aqueous phase to a fresh polypropylene tube. Be careful to avoid the protein interface, if any.

Note: If a band of protein persists between the aqueous and organic phases, it would be wise to repeat the organic extraction (steps 15–17) until all of the protein has been removed.

18. Perform one final organic extraction by the addition of an equal volume of chloroform alone or a mixture of chloroform:isoamyl alcohol (24:1). Centrifuge the sample at $1000 \times g$ for 3 min at 4° .
19. Transfer the aqueous phase to an RNase-free Corex⁶ glass tube that has been prechilled on ice. Add 3 vol 100% ethanol. Cover each tube with Parafilm. If it does not appear that the ethanol and aqueous solution have mixed, then *carefully* and *gently* vortex the tube for a few seconds. *Do not* invert the tube to mix.
20. Place each tube at -20° and store overnight in order to maximize the recovery of RNA.
21. Place Corex tubes into suitable rubber adapters in a precooled rotor and centrifuge at $9000 \times g$ for 30 min at 4° .
22. At the conclusion of the run, carefully decant and discard the supernatant, taking care not to lose any of the precipitated material on the side or bottom of the tube.
23. Wash RNA with 75% ethanol. Use 5 ml to wash RNA that was precipitated in a 15 ml Corex tube or 10 ml to wash RNA that was precipitated in a 30 ml Corex tube.
24. Vortex the tube(s) gently, and then centrifuge at $9000 \times g$ for 15 min at 4° .
25. Repeat steps 22–24 twice, for a total of three washes.
26. Wash RNA once with 5 ml of 95% ethanol. Vortex gently, and then centrifuge at $9000 \times g$ for 10 min at 4° .

Note: Washing the RNA in 100% ethanol will expedite the drying process and favor maximum recovery of RNA. The very high-quality RNA that will result should perform extremely well in any of a variety of downstream applications.

27. Allow the RNA in the tube to air dry, but take care to avoid drying out the sample completely. Having washed the precipitate with 100% ethanol, the sample should be dry within 5 min. The drying process can be further accelerated by placing the tube on a downward angle in a 37° incubator for no more than 2–3 min. Be sure that the tube is uncovered so residual ethanol can drain out; also ensure that the opening of the tube is resting on a stack of 3–4 Kimwipes, and not directly on the metal inside the incubator.
28. Resuspend the RNA in the smallest possible volume of nuclease-free H_2O or TE buffer (10 mM Tris, pH 7.5; 0.1 mM EDTA). Depending on the starting mass of the

⁶Be sure to use only those high-strength tubes that can withstand the required RCF. If Corex tubes are not available, an equivalent product can be purchased from Kimble-Kontes. Use catalogue number 45500-15 for 15 ml tubes and 45500-30 for 30 ml tubes. As with Corex tubes, the Kimble high-strength tubes must be placed in a suitable rubber adapter for centrifugation.

tissue completely dissolving the RNA may require a volume of as much as 500–1000 μ l or more.

29. Remove an aliquot of the sample for spectrophotometric determination of the concentration of the RNA (Chapter 6). Next, run an aliquot on a denaturing agarose gel (Chapter 9) to assess sample quality.

Note: Smearing on the gel long the entire length of the lane may be an indicator that the RNA has not completely dissolved (Fig. 3.5).

30. Distribute RNA into convenient aliquots for storage at -80° . RNA can be stored in hydrated form or can be reprecipitated for long-term storage.

Protocol: RNA isolation from lipid-enriched samples

The isolation of RNA from adipose tissue is a simple matter, being accomplished most easily through the use of guanidinium buffers to disrupt the tissue architecture. Perhaps the most cumbersome issue related to this method is removal of the preponderance of lipids stored in adipose- and other types of tissue, including brain and breast tissue. In general, cells or tissues are homogenized in guanidinium buffer (or LiCl-urea homogenization buffer), followed by a brief, low speed centrifugation during which the lipid floats to the top of the tube, from which it may then be aspirated away. This low-speed centrifugation is repeated as needed. Alternatively, one may begin with the addition of chloroform directly to a guanidinium lysate, in order to draw the lipids into the organic phase, which is then removed. RNA isolation then continues as usual with the addition of acidic phenol and chloroform, followed by the purification of the RNA. While the following procedure (a modification of Chomczynski and Sacchi, 1987) assumes the investigator will be preparing in-house reagents, one may confidently substitute commercial RNA isolation reagents such as Trizol or TRI Reagent.

1. Wear gloves!
2. Mince freshly isolated tissue (up to 100 mg) on ice. For each 0.1 g of tissue add 1 ml of solution D:
 - 4 M guanidinium thiocyanate
 - 25 mM sodium citrate, pH 7.0
 - 0.5% sarcosyl
 - 100 mM 2-mercaptoethanol (*add just prior to use*)and homogenize at room temperature. Add an equal volume of chloroform and then pulse centrifuge to separate the phases. Remove and discard the chloroform (lower) phase. Add a fresh volume of chloroform, pulse centrifuge to separate the phases, and then transfer the upper aqueous phase to a fresh polypropylene tube.
3. For each 1 ml of solution D lysis buffer used in step 2, add:
 - 0.1 ml 2 M sodium acetate, pH 5.2
 - 1.0 ml water-saturated phenol (molecular biology grade)
 - 0.2 ml chloroform:isoamyl alcohol (49:1)Cap tube and mix carefully and thoroughly by inversion following the addition of each reagent and invert vigorously for an additional 30 s after all reagents have been added.
4. Cool sample on ice for 15 min.
5. Centrifuge to separate the phases; for microfuge tubes use $12,000 \times g$ at 4° , if possible.

6. Transfer aqueous phase (containing RNA) to a fresh tube and mix with 0.75 volume of ice-cold isopropanol. Store at -20° for at least 1 h to precipitate RNA.
Note: For larger scale preparations, Corex glass tubes, placed in the correct adapters, can be used for precipitation and recovery of RNA.
7. Collect precipitate by centrifugation at $9000 \times g$ for 10 min at 4° . Carefully decant and discard supernatant. **CAUTION:** Do not exceed the maximum g-force for any of the tubes used in this protocol.
8. Completely dissolve RNA pellet in $300\mu\text{l}$ of solution D (see step 2) and then transfer to an RNase-free 1.7 ml microfuge tube.
9. Reprecipitate the RNA by the addition of 0.75 volume of ice-cold isopropanol and store at -20° for 1 h.
10. Collect precipitate at top speed in a microcentrifuge for 10 min at 4° . Carefully decant and discard supernatant.
11. Wash pellet 3–4 times in $500\mu\text{l}$ 70% ethanol. If the RNA does not dislodge during these washes, there is no need to re-centrifuge. Allow tubes to air-dry to remove residual ethanol. Optional: wash sample once more with 95% ethanol to accelerate the drying process.
12. Dissolve the RNA in the smallest possible volume of TE buffer or nuclease-free H_2O . Incubation at 65° for 10 min may facilitate solubilization, though this is unnecessary if the RNA did not dry out completely following the ethanol washes. Store RNA as an ethanol precipitate until ready to use. Following determination of concentration, store the RNA in suitable aliquots at -80° . Avoid repeated freezing and thawing.

Purification of polysome-engaged mRNA

mRNA is an excellent parameter of gene expression but it is not the only one. Thorough assessment of regulation of this aspect of the cellular biochemistry is multifaceted. Standard RNA isolation techniques, even when coupled to PCR, reveal information about the steady-state abundance of certain RNAs in the cell at the moment of cell lysis, yet reveal nothing at all about the translational fate of the transcripts of experimental interest. Recall that the gene expression is also controlled at the translational and posttranslational levels.

Clearer definition of the translational aspect of gene expression may be gained by collecting and analyzing the mRNA fraction that has engaged the translational machinery. The polysome fraction of the cell (all mRNAs engaged by ribosomes) is a fairly accurate indicator of the proportion of the mRNA mass that has actually advanced to the translational level along the gene expression pathway. In the cell, some polysomes are associated with the endoplasmic reticulum (presumably synthesizing secreted proteins, mitochondrial proteins, and proteins embedded in the membrane) while others remain as free polysomes, which are believed to synthesize proteins that will remain in the cytoplasm or move into the nucleus. Thus, the isolation of polysome-engaged mRNA is used to profile gene expression at the translational level and may well be a more accurate indicator of both efficiency of translation initiation as well as the phenotype identity of the cells under investigation.

When contemplating gene expression it is important to remember that the abundance of mRNA does not always correlate well with the abundance of

its cognate protein. Succinctly, polysomal RNA purified from cell culture or from tissue can be analyzed to determine which mRNAs are being translated and with what efficiency. This type of data can be useful for understanding the dynamic foundation of the proteome because functional proteins are observed only after experiencing post-translational modification(s).

Methods describing the immunological precipitation of specific polysomes first appeared in the literature in the early 1970s (Palacios *et al.*, 1972; Schechter, 1973; Palmiter, 1974). Within a few years, the methodology was improved upon by restructuring the mechanics of the technique to the affinity column format (Schultz *et al.*, 1977). Among the more noteworthy examples of the efficacy of this method as an RNA enrichment technique was the isolation of the mRNA encoding the heavy chain of the human HLA-DR histocompatibility antigen using epitope-specific monoclonal antibodies (Shapiro and Young, 1981; Korman *et al.*, 1982). Shortly thereafter, polyvalent sera was used successfully for the isolation of mRNAs encoding cystathione β -synthase from rat liver (Kraus and Rosenberg, 1982) and the low-density-lipoprotein receptor from bovine adrenal cells (Russell *et al.*, 1983). In general, the use of antibodies for polysome immunopurification was difficult to perform and offered only limited success because many of the well-characterized antibodies available at the time recognize antigens that were glycosylated or had otherwise matured, a characteristic not associated with nascent proteins.

Approaches such as these are no longer favored as enrichment techniques for specific mRNAs because of the widespread use of PCR. Newer applications involve the use of the cellular polysome fraction as method for characterizing overall translational control rather than for the isolation of one or more specific messages by antibody recognition, and several noteworthy examples have recently been published (Kash *et al.*, 2002; Del Prete *et al.*, 2007). Expression profiling analysis in this manner is a rather accurate indicator of translational efficiency because of the high correlation between formation of the translation apparatus and the synthesis of the mRNA-encoded primary polypeptide. The availability of on-line genome databases is also extremely helpful for the design of primers, reducing the ambiguities that might otherwise be associated with the targeting of specific genes. Now, a standard approach is the precipitation of all polysomes to pull down that subpopulation of mRNA that is actively undergoing translation. Then, free from the ribosomes and ancillary translation factors, purified polysomal mRNA can be reverse transcribed to synthesize probes for microarray analysis (Chapter 21), assayed by PCR with gene-specific primers, enriched using universal primers to support the construction of a cDNA library, or sequenced directly. The method which follows is a variant of the magnesium precipitation procedure described by Palmiter (1974). It has the advantage of not requiring ultracentrifugation and is compatible with tissue and cells grown *in vitro*. From an RNase control perspective, the heparin is a competitive inhibitor of RNase as is magnesium at the elevated concentration used in the precipitation and polysome buffers.

Protocol: isolation of polysomal mRNA

1. Wear gloves! Be sure to observe RNase-free technique throughout this procedure and perform all centrifugations at 4° unless otherwise noted.
- 2a. *Tissue samples:* The starting mass will, of course, be variable from one sample to the next and will also be dependent on tissue availability. For each mg tissue add 100 μ l of grinding buffer (25 mM NaCl; 5 mM MgCl₂; 25 mM Tris, pH 7.5; 1.5% Triton X-100; 0.5 mg/ml heparin). Homogenize the tissue on ice in a suitably-sized Dounce homogenizer with a loosely-fitting pestle. Add small volumes of additional grinding buffer, as needed. Go to step 3.

Note: Tissue can be flash-frozen in liquid N₂ and then pulverized with a nuclease-free mortar and pestle prior to the addition of the grinding buffer.

- 2b. *Cell culture:* Harvest the cells as usual (see Appendix J). In this protocol in particular it is important to wash the cell monolayer with ice cold PBS and to use spent cell culture medium to arrest the activity of the trypsin. For each 10⁶ cells add 100 μ l of grinding buffer (25 mM NaCl; 5 mM MgCl₂; 25 mM Tris, pH 7.5; 1.5% Triton X-100; 0.5 mg/ml heparin).

Note 1: RNasin, at a final concentration of 200 U/ml can be used in addition to heparin as an RNase inhibitor. If heparin is used, it will be necessary to remove all traces of it (below) if the RNA is to be reverse transcribed.

Note 2: In order to maintain polysome integrity throughout the isolation procedure, the cell culture growth medium can be supplemented with up to 50 μ g/ml cycloheximide for no more than 15 min just prior to harvesting. The idea is to prevent polysome dissociation prior to mRNA capture and this treatment will, in fact, maximize polysome recovery.

3. Transfer the homogenate into a suitable RNase-free centrifuge tube. For larger volumes, use a silanized Corex glass tube, and don't forget to place the tube(s) into a suitable rubber adapter. For smaller volumes, use an RNase-free microfuge tube. In either case, centrifuge for 5 min at 10,000 \times g in a prechilled rotor (be sure not to exceed the g-force rating for the tube).
4. While the sample is being centrifuged, prechill an appropriate number of sterile conical 50 ml tubes on ice.
5. At the conclusion of the centrifugation in step 3, carefully decant the supernatant into the prechilled tubes. Depending on the original volume of the homogenate, it may be necessary to split the supernatant into two tubes in order to accommodate the subsequent required reagents.
6. Add 1.2 vol of precipitation buffer (25 mM NaCl; 200 mM MgCl₂; 25 mM Tris, pH 7.5; 1.5% Triton X; 0.5 mg/ml heparin or 200 U/ml RNasin). Incubate on ice for 2 h.
7. At the conclusion of the incubation period, add 3.5 ml polysome buffer (25 mM NaCl; 200 mM MgCl₂; 25 mM Tris, pH 7.5; 1 M sucrose) into sterile, silanized 15 ml Corex glass tube(s). Then, carefully overlay 10 ml of the mixture from step 6, above.

Note: See Appendix I for silanization procedure.

8. Centrifuge for 15 min at 10,000 \times g at 4° in a prechilled rotor (swinging bucket rotor is best).
9. *Carefully* decant and discard the supernatant, taking care not to lose the pellet. If the pellet appears loose, then remove the supernatant by gentle aspiration.
10. Resuspend the polysome pellet in 1 ml of room temperature HEPES-SDS buffer (0.1 M NaOAc; 10 mM HEPES, pH 7.5; 0.1% SDS). More of this buffer can be added to completely redissolve the pellet as needed.

11. If the volume of the resuspended pellet from step 10 is 1 ml or less, transfer all of it to a 2.2 ml microfuge tube. If the volume of the resuspended pellet from step 10 is more than 1 ml, it will be necessary to either split the volume into two microfuge tubes.
 12. Add an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) to each tube. Secure the cap and then mix carefully. Centrifuge for 5 min at $10,000 \times g$ to separate the phases.
- Note: If larger tubes are used for the extraction, be sure not to exceed the maximum recommended RCF for that tube.*
13. Transfer the aqueous (upper) phase to a fresh tube, and then repeat step 12.
 14. Transfer the aqueous (upper) phase to a fresh tube and then add an equal volume of chloroform or chloroform:isoamyl alcohol (24:1). Centrifuge for 30 s at $10,000 \times g$ to separate the phases.
 15. Transfer the aqueous material to a larger volume centrifuge tube that can withstand $10,000 \times g$, such as a (silanized) Corex glass tube. If any samples were split into multiple tubes in step 11, they can now be recombined.
 16. Add 0.08 vol 3 M NaOAc, pH 5.2 and 2.5 vol 100% ethanol. Store at -20° for a minimum of 2 h to precipitate the RNA; an overnight incubation at -20° will maximize RNA recovery.
 17. Collect the RNA by centrifugation at $10,000 \times g$ for 20 min at 4° .
 18. Carefully decant the ethanol, taking care not to lose the precipitated RNA.
 19. Wash to RNA twice with 75% ethanol.
 20. Wash the RNA once with 95% ethanol to facilitate rapid air-drying. Do not allow the RNA to dry out.
 21. Resuspend the RNA in the smallest possible volume of nuclease-free H_2O . Gently vortexing the sample will help dissolve the material in the tube. If necessary, heat the tube in a 65° water bath for no more than 2 min to help dissolve the sample and then vortex gently again.
 22. Optional: precipitate the RNA in 2 M LiCl overnight at 4° to remove all traces of heparin. Collect the precipitate at $10,000 \times g$ for 15 min at 4° in a suitable centrifuge tube. Wash pellet twice with 70% ethanol to wash away residual salt and then wash once with 95% EtOH to expedite air-drying. Dissolve the RNA in a suitable aliquot of sterile water or TE buffer (pH 7.5).
 23. Determine the concentration and purity of the RNA as described in Chapter 6. Aliquots of this RNA should be stored at -80° until ready for use.

Collecting samples in the field

A very important compound that has revolutionized the way samples are stored for RNA analysis at a later time, date, or location is RNAlater[®] (Ambion). This product is a proprietary aqueous formulation that is well known in RNA circles for its ability to prevent deterioration of RNA in tissues and organs, until such time as it is convenient for the investigator to process the biological material and recover high-quality RNA for any of several uses. RNAlater quickly permeates cells and tissue samples thereby allowing safe storage of tissue for up to a week at room temperature, for as long as a month at 4° , and indefinitely at -20° or colder (Ambion product specification sheet).

The chemical attributes of *RNAlater* offer several strategic advantages with respect to the collection of samples and subsequent processing, especially when working out in the field. First, it is generally not practical to transport even a small dewar of liquid nitrogen for the purpose of flash freezing, especially if the investigator is planning to be in the field for several days. Second, RNA in a tissue sample is immediately stabilized. As such, tissue samples and their RNA contained therein are stabilized without risk of degradation. Third, *RNAlater* is effective for the storage of Gram-positive and Gram-negative bacteria, as well as tissue culture cells, thereby giving the investigator much greater latitude in terms of when the RNA must actually be isolated from the cells. Fourth, the use of *RNAlater* is compatible with most RNA isolation procedures associated with various downstream applications. Finally, because of its non-toxic composition, *RNAlater* can be disposed of in the sink, accompanied by thorough flushing with water.

RNA “Clean-Up” methods

No matter which procedure is used to isolate RNA from fresh-, frozen-, or fixed archival tissue, the RNA is often contaminated with small amounts of residual protein or compounds used as part of the isolation procedure. These contaminants can have a negative impact on the utility of the RNA in most popular applications. If the electrophoretic appearance of newly isolated RNA is good but the RNA does not support conversion into cDNA or otherwise function in any type of hybridization assay, then it may be necessary to clean up the RNA. This can be performed in any number of ways. For example, one may increase the volume of the dissolved RNA a bit with sterile water or TE buffer and then extract it with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), as described in this chapter and in Chapter 2. It is definitely wise to keep the combined volume of RNA and the organic material under 1500 μ l so that the extraction can be performed in a microfuge tube. Upon centrifugation, the appearance of a white band between the aqueous- and organic phases is a sure indicator that all of the protein was not removed. The extraction should be repeated until all of the protein is removed, as demonstrated by the absence of a band between the upper and lower phases. Then, the aqueous phase should be extracted once with an equal volume chloroform *or* chloroform:isoamyl alcohol (24:1) to remove all traces of phenol that could oxidize into very damaging quinones. Next, the RNA can be salt and alcohol precipitated and quality control-assayed again.

Alternatively, there are a number of silica-based spin columns that bind RNA under high-salt chaotropic conditions, as described previously. These products, available from nearly all vendors of molecular biology wares, allow the RNA to be sequestered from all of the other chemicals and compounds in a tainted sample. Purified RNA is then eluted from the column in low salt buffer or in sterile water. This approach is rapid and, although it often utilizes guanidinium thiocyanate as the chaotropic agent, it does not involve the use of any organic solvents. In this lab, we have successfully cleaned up suspect RNA

samples by redissolving them in the RNA lysis buffer that accompanies spin column RNA isolation kits. This sample is then processed as if working with a whole cell lysate. The results of this type of cleanup approach are remarkable.

RNA clean-up can also mean the removal of contaminating genomic or mitochondrial DNA from the sample. Recall that the more aggressive methods directed toward RNA isolation disrupt all organelles, thereby liberating DNA into the homogenate. It is standard operating procedure in most labs to treat RNA isolated from tissue with DNase I, as described in Appendix F. In so doing, the investigator will not need to worry about competition between cDNA and genomic DNA during PCR. Nor will contaminating DNA produce unwelcome bands during blot analysis.

Finally, it is important to recognize the fact that certain tissues are characterized by excessive quantities of natural cell products that may copurify with the RNA. Noteworthy examples include the copurification of glycogen from liver tissue and the copurification of lipid from adipose tissue. In the case of the former, the addition of 3 vol of 4M sodium acetate, pH 7.0 (*not* pH 5.2) followed by incubation overnight at -20° will precipitate the RNA while leaving the glycogen behind in solution (Sambrook *et al.*, 1989). The RNA is covered by centrifugation at a minimum of $8000 \times g$ (depends on the centrifuge tube) for 15 min at 4° . Alternatively, adding 4M LiCl to a tube containing precipitated RNA, and breaking up the pellet, will move the contaminating glycogen away from the RNA (Puissant and Houdebine, 1990); the RNA is then recovered by centrifugation for 10 min at $3000\text{--}5000 \times g$, depending on the tube. In the case of the latter, strategies for dealing with excessive quantities of lipid are described above.

The take home lesson: be attentive to the details of the various isolation procedures in order to generate the highest quality RNA possible. It is wise to minimize the number of manipulations to which the RNA must be subjected so as to favor maximum stability of the sample while minimizing the risk of accidental introduction of RNase. RNA does not withstand repeated precipitations that might be needed if the sample was not properly purified the first time.

Troubleshooting RNA isolation from tissue

When the isolation of RNA from tissue samples is marginal or fails altogether, the reason may be due to any or all of the following:

1. *Incomplete tissue disruption.* This is an especially important concern because, unlike the lysis of cells propagated in culture, the three-dimensional architecture of the tissue is formidable due to the presence of resilient proteins in the extracellular matrix as well as myriad intercellular junctions. The lysis conditions may need to be modified, and homogenization by mechanical means, rather than chemical disruption, may be necessary. Some tissues are notoriously difficult to break up.
2. *Overly zealous homogenization.* The forces associated with tissue disruption are usually needed to effect liberation of cellular components. In practice, short bursts with an electronic homogenizer or fewer, rather than more, strokes with a Dounce

homogenizer will usually yield RNA of the best possible quality. Limit the use of the homogenization device to only what is necessary to do the job.

3. *Failure to control RNase activity.* Some animal organs and tissues contain abundant levels of RNase. Degradation of RNA can begin after removal of the tissue sample from its biological source if not frozen immediately or otherwise stabilized. Modify lysis buffers with one or more non-specific inhibitors of RNase or use buffers that are highly chaotropic.
4. *Incomplete removal of protein.* The methods described herein and elsewhere for RNA isolation from tissue, when applied correctly, usually produce high-purity RNA as assessed by A_{260} and A_{280} measurements. When the A_{260}/A_{280} ratio is below 1.8 (a ratio of 2.0 is ideal), one should consider dissolving the RNA in a fresh aliquot of lysis buffer and then proceed to purify it again as if starting with a sample of cells grown in culture (re-homogenization not necessary). Paying careful attention to the lab protocol and the recommended volumes will generally preclude this difficulty.
5. *Too much tissue in too little buffer.* A common mistake among those investigators working with tissue for the first time is the failure to use enough lysis buffer based on the mass of the tissue sample. The rule of thumb when working with cells produced in culture is to keep the sample as concentrated as possible in order to facilitate RNA precipitation at the end of the procedure. It is easy to underestimate the number of cells in a milligram or more of tissue, and using too little lysis buffer results in incomplete separation of macromolecules. Even when the proportion of lysis buffer to tissue is correct, investigators often fail to use a suitable amount of sterile water or TE- or other buffer to dissolve the purified RNA. This makes accurate spectrophotometric measurement difficult and electrophoretic profiling all but impossible. Remember that many tissue samples often yield very large quantities of RNA and the volumes of the lysis buffer and buffer used to dissolve the purified RNA need to be scaled appropriately.

References

- Auffray, C. and F. Rougeon. (1980). Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* **107**, 303–314.
- Chomczynski, P. and N. Sacchi. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analyt. Biochem.* **162**, 156–159.
- Chung, J.-Y., T. Braunschweig and S.M. Hewitt. (2006). Optimization of recovery of RNA from formalin-fixed, paraffin-embedded tissue. *Diagnostic Mol. Pathol.* **15**, 229–236.
- Coombs, N.J., A.C. Gough and J.N. Primrose. (1999). Optimization of RNA and RNA extraction from archival formalin-fixed tissue. *Nucleic Acids Res.* **27**(16), e12–e14.
- Del Prete, M.J., R. Vernal, H. Dolznig, E.W. Mullner and J.A. Garcia-Sanz. (2007). Isolation of polysome-bound mRNA from solid tissues amenable for RT-PCR and profiling experiments. *RNA* **13**, 1–8.
- Fend, F., M.R. Emmert-Buck, R. Chuaqui, K. Cole, J. Lee, L.A. Liotta and M. Raffeld. (1999). Immuno-LCM: laser capture microdissection of immunostained frozen sections for mRNA analysis. *Am. J. Pathol.* **154**(1), 61–66.
- Kash, J.C., D.M. Cunningham, M.W. Smit, Y. Park, D. Frtiz, J. Wilusz and M.G. Katze. (2002). Selective translation of eukaryotic mRNAs: functional molecular analysis of GRSF-1, a positive regulator of influenza virus protein synthesis. *J. Virol.* **76**, 10417–10426.

- Kato, N., S. Pfeifer-Ohlsson, M. Kato, E. Larsson, J. Rydnert, R. Ohlsson and M. Cohen. (1987). Tissue-specific expression of human provirus ERV3 mRNA in human placenta: two of the three ERV3 mRNAs contain human cellular sequences. *J. Virol.* 61(7), 2182–2191.
- Kuecker, S.J., L. Jin, E. Kulig, G.L. Oudraogo, P.C. Roche and R.V. Lloyd. (1999). Analysis of PRL, PRL-R, TGF-R11 gene expression in normal and neoplastic breast tissues after laser capture microdissection. *Appl. Immunohistochem. Mol. Morphol.* 7(3), 193–200.
- Korman, A.J., P.J. Knudsen, J.F. Kaufman and J.L. Strominger. (1982). cDNA clones for the heavy chain of HLA-DR antigen obtained after immunopurification of polysomes by monoclonal antibody. *Proc. Natl. Acad. Sci.* 70, 1844–1848.
- Kraus, J.P. and L.E. Rosenberg. (1982). Purification of low-abundance messengers RNAs from rat liver by polysome immunoadsorption. *Proc. Natl. Acad. Sci.* 79, 4015–4019.
- Masuda, N., T. Ohnishi, S. Kawamoto, M. Monden and K. Okubo. (1999). Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res.* 27(22), 4436–4443.
- Mikhitarian, K., S. Reott, L. Hoover, A. Allen, D.J. Cole, W.E. Gillanders and M. Mitas. (2004). Enhanced detection of RNA from paraffin-embedded tissue using a panel of truncated gene-specific primers for reverse transcription. *BioTechniques* 36(3), 474–478.
- Palacios, R., R.D. Palmiter and R.T. Schimke. (1972). Identification and isolation of ovalbumin-synthesizing polysomes. I. Specific binding of ^{125}I -anti-ovalbumin to polysomes. *J. Biol. Chem.* 247, 2316–2321.
- Palmiter, R.D. (1974). Magnesium precipitation of ribonucleoprotein complexes. Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* 13(7), 3606–3615.
- Puissant, C. and L. Houdebine. (1990). An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *BioTechniques* 8(2), 148–149.
- Russell, D.W., T. Yamamoto, W.J. Schneider, C.J. Slaughter, M.S. Brown and J.L. Goldstein. (1983). cDNA cloning of the bovine low density lipoprotein receptor: feedback regulation of a receptor mRNA. *Proc. Natl. Acad. Sci.* 80, 7501–7505.
- Sambrook, J., E.F. Fritsch and T. Maniatis. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY p. 7.22.
- Schechter, I. (1973). Biologically and chemically pure mRNA coding for a mouse immunoglobulin L-chain prepared with the aid of antibodies and immobilized oligothymidine. *Proc. Natl. Acad. Sci.* 70, 2256–2260.
- Schultz, G., S. Kieval, B. Groner, A.E. Sippel, D.T. Kurtz and P. Feigelson. (1977). Isolation of specific messenger RNA by adsorption of polysomes to matrix-bound antibody. *Nucleic Acids Res.* 4, 71–84.
- Shapiro, S.Z. and J.R. Young. (1981). An immunochemical method for mRNA purification. Application to messenger RNA encoding trypanosome variable surface antigen. *J. Biol. Chem.* 256, 1495–1498.
- Stanta, G. and C. Schneider. (1991). RNA extracted from paraffin-embedded human tissues is amenable to analysis by PCR amplification. *BioTechniques* 11(3), 304–308.
- Suarez-Quian, C.A., S.R. Goldstein, T. Pohida, P.D. Smith, J.I. Peterson, E. Wellner, M. Ghany and R.F. Bonner. (1999). Laser capture microdissection of single cells from complex tissues. *BioTechniques* 26(2), 328–335.

This page intentionally left blank

4 Going Green: RNA and the Molecular Biology of Plants

Rationale	105	Protocol: RNA isolation from fruit	113
RNA isolation and the peculiarities of plants	105	Strategies for RNA isolation from other plant tissues	115
Types of RNA produced in plant cells	109	Troubleshooting RNA isolation from plant tissue	116
Protocol: RNA isolation from leaf	110	References and suggested reading	117
Protocol: RNA isolation from bark	112		

Rationale

As eukaryotic organisms, plants exhibit complex gene expression pathways and regulatory controls that are often similar to those observed in animals. Consequently, techniques that are commonly used to assay gene expression in animal cells may be used to assay the same in *Plantae*. The increased awareness and interest in agricultural biotechnology and applications thereof have created a need for expedient isolation of high-quality RNA from vastly different plant tissues, such as leaf, root, bark, floral tissue, fruit, and cotyledon. The fact that plants are unable to move when local conditions become unfavorable makes RNA profiling of plant tissue particularly fascinating, and much is to be gained by understanding the molecular biology of plants. Several contemporary lines of research include identifying the genes responsible for cold-hardiness, discerning plant molecular responses to the biotic (e.g. pathogen) and abiotic (e.g. drought) stress, and identifying genes that may lead to the cost-effective synthesis of biofuels to name but a few examples. The good news is that plant RNA seems to be remarkably stable when compared to the animal cell counterpart, meaning that there are often much greater opportunities for long-term storage and analysis.

RNA isolation and the peculiarities of plants

The most obvious ultrastructural features of plants cells that distinguish them from animal cells are (1) the presence of a cell wall; (2) the presence of a large

fluid-filled vacuole; and (3) the presence of chloroplasts. While it is true that RNA is chemically the same type of molecule regardless of its biological source, the rather fibrous nature of plant tissue often requires far more aggressive measures to disrupt cells than are required for an equal starting mass of animal tissue. A number of isolation procedures have been described, and kits developed, to deal with the unique qualities of individual plant species, tissues, and organs. There is no standard protocol for RNA isolation from all plant species. Examples of strategies used to isolate RNA from various plant structures are presented here; procedures best suited for one's own research will have to be determined empirically.

The plant cell wall is obviously structural in nature, contributing to defense at both the cell and organism levels and regulating the distribution of water, minerals, and growth-associated compounds. From an experimental perspective, the plant cell wall represents a challenge for getting macromolecules out of cells as well as for getting things into cells, such as new gene constructions or siRNA molecules for gene silencing (Chapter 23). The plant cell wall is quite a bit thicker than the cell membrane that it surrounds and this rigid structure is the primary reason why plants are sessile organisms.

The major components of what is known as the primary cell wall are cellulose microfibrils, hemicellulose, pectins, and glycans which have a cross-linking role in the architecture of this remarkable structure. As the cell and the wall surrounding it mature, the formation of what is known as the secondary cell wall is commonly observed. While there are (often extensive) biochemical similarities between the primary and secondary cell walls, compounds that are most often associated with the secondary cell wall include lignin, which is a composite of aromatic alcohol polymers, as well as cutin, suberin, and other materials. In a manner very much like the extracellular matrix in animal cells, the plant cell wall tethers cells together resulting in tissue formation. From an RNA isolation perspective, the cell wall is yet another barrier between the investigator and the transcripts.

Mature plant cells characteristically have a rather large fluid-filled central vacuole in the cytoplasm. Bound by a dynamic membrane known as the tonoplast, the vacuole is filled with a variety of dissolved organic and inorganic compounds, including compounds that are distasteful to insects and other plant-eating animals. Among its many functions, the vacuole maintains turgor pressure, meaning the water in the vacuole exerts an outward pressure against the cell wall from the within the cell. Turgor pressure is what helps plants remain upright and contributes to the characteristic shape of plant cells. Loss of turgor pressure results in wilting, and the cell wall also prevents plant cells from lysing should a plant find itself in a hypotonic environment. From a molecular biology perspective, high cellular water content often translates to a low nucleic acid concentration.

The aerial tissues in plants contain immense numbers of round-to-oval-shaped chloroplasts. Like mitochondria, chloroplasts have their own genome and associated transcripts. During RNA isolation procedures the chloroplasts rupture,

thereby liberating very large amounts of chloroplast mRNA. These transcripts are highly abundant and characteristically appear as very distinct bands between and below the ribosomal subunits (Fig. 4.1). These bands are often unsettling to people who are accustomed to seeing only the predominant rRNA species upon electrophoretic examination. One should consider the appearance of the chloroplast transcripts to be an indicator of the integrity of the sample, meaning that downstream applications such as RT-PCR can be approached with confidence in the template material.

Among the first considerations when planning to isolate plant RNA is the method for dismantling or otherwise breaking through the formidable cell wall. Methods that have been used to meet this challenge include pulverizing plant tissue that has been frozen in liquid nitrogen (the predominant method), the use of an electronic homogenizer or a BeadBeater (discussed in Chapter 3), or a laboratory blender or similar apparatus. While soft tissues can often be homogenized without first freezing it, the achievable level of disruption is often suboptimal, resulting in a lower RNA yield. After pulverizing the tissue, a short centrifugation step ($10,000 \times g$) is almost always needed for removal insoluble plant cell- and tissue debris. From this point forward, the processing of the cell homogenate varies depending on the type of plant tissue starting

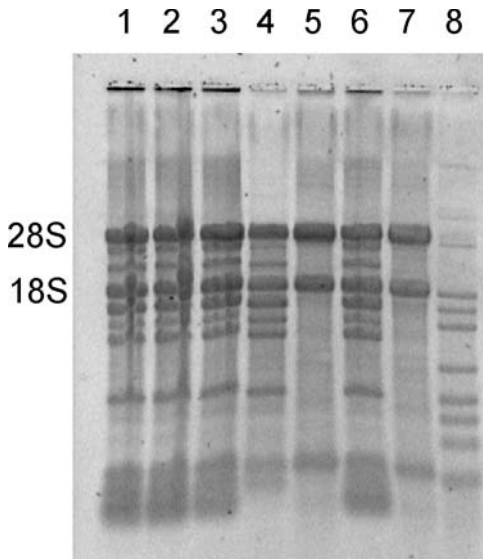


Figure 4.1 Electrophoretic profile of plant RNA from green tissue. Note the abundant chloroplast transcripts above, between, and below the 28S and 18S rRNAs. These transcripts, as well as the well-formed ribosomal species, demonstrate the integrity of the sample. While the stability of certain plant transcripts is legendary, it is wise to run a gel just to be sure that the RNA is intact, especially if more than six months have elapsed since the RNA was last examined. Electrophoresis of RNA is described in detail in Chapter 9.

material, the intended use of the RNA, and the preference(s) of the investigator. In some instances, for example, the resulting supernatant is extracted with phenol, including a heating step, to remove cellular proteins, and then chloroform-extracted to remove residual phenol. Bringing the sample to a final concentration of 2–3 M LiCl selectively drives the precipitation of RNA, allowing it to be sequestered from contaminating DNA and other impurities that would otherwise copurify with the RNA. The recovered RNA can then be precipitated again with sodium acetate and ethanol, which will allow the RNA to be cleaned up nicely. This approach to RNA isolation is remarkably similar to the phenol-based methods that have been used for RNA isolation from animal cells, with the exception of the initial pulverization of flash-frozen tissue, and is considered by most investigators to be the “classical method” by which to approach plant RNA isolation.

A somewhat variant approach to RNA isolation is through the use of a guanidinium buffer, usually guanidinium thiocyanate¹, as a phenol alternative. Needless to say, great care must be taken at all levels to control RNase during plant RNA isolation procedures. The chaotropic nature of guanidinium reagents (e.g. Trizol) characteristically inactivates nucleases on contact and tends to move proteins away from nucleic acids by solubilizing them. As always, the success of these methods depends upon using the proper ratio of lysis buffer to starting mass. Keeping the sample fairly concentrated at all times is also a valuable strategy and, logistically speaking, it is much easier to work at the microfuge tube scale than having to use larger (e.g. Oak Ridge and Corex glass) tubes and centrifuges. Recently, an alternative isolation method has been reported (Bligin *et al.*, 2009) in which a combination of the chaotropic agents ammonium thiocyanate and guanidine thiocyanate are used in the presence of acidic phenol. This method has the advantage of removing more organic compounds and rendering improved A_{260}/A_{280} ratios than previously published methods for the recovery of RNA from plants.

Some of the other time-honored methods for RNA isolation from plant tissue feature the detergent CTAB (Cetyltrimethylammonium bromide), which quite efficiently promotes the separation of both DNA and RNA from polysaccharides. Other more traditional methods for plant RNA recovery likewise favor aggressive disruption in the presence of high concentrations of SDS. Regardless of the method of lysis and subsequent processing, one of the major concerns associated with the isolation of plant RNA is carbohydrate contamination, which should be suspected if the A_{260}/A_{230} ratio is low. As with glycogen contamination in RNA isolated from animal liver tissue, carbohydrates contaminating plant RNA preps can be removed by precipitating the RNA in LiCl. Briefly, RNA can be precipitated directly from an aqueous solution by the addition of LiCl to a final concentration of 3 M LiCl and then incubating overnight at -20° . The RNA can then be recovered by microcentrifuging the

¹The use of guanidine hydrochloride for this application has also been reported (Jaiprakash *et al.*, 2003).

sample(s) for 10 min at $12,000 \times g$, preferably at 4° . Fortunately, several of the products designed for plant RNA isolation are likewise designed to minimize these contaminants in the final RNA prep.

Types of RNA produced in plant cells

RNA molecules in plant cells arise from one of three cellular compartments, namely the nucleus, the mitochondria, and the chloroplasts. In all eukaryotic organisms, there are (at least) four known nuclear polymerases, namely RNA polymerase I, RNA polymerase II, RNA polymerase III, and RNA polymerase IV, as reviewed in Chapter 1. The predominant mature products of nuclear transcripts are, ranked by mass, rRNA, tRNA, and mRNA (a mature, spliced form of hnRNA). Nuclear RNA is transported from the nucleus and into the cytoplasm through the formation of ribonucleoprotein (RNP) particles by transient association of transcripts with various RNA binding proteins. RNA polymerase IV in plants (Onodera *et al.*, 2005) does not appear to overlap in function with RNA polymerases I, II, or III. Rather, RNA polymerase IV in plants appears to have a role promoting methylation-associated higher-order heterochromatin formation (Kanno *et al.*, 2005). Finally, some RNA molecules are transported from the cytoplasm into the mitochondria, such as certain tRNA species that are not encoded in mtDNA but are required to support protein synthesis in the mitochondria.

Angiosperms also exhibit a unique group of nuclear RNA polymerases that strongly resemble the very well-characterized T7 and T3 RNA polymerases from bacteriophage. Some members of this so-called RpoT gene family can be found both in monocots and dicots. In *Arabidopsis*, RpoT1 is responsible for the transcription of mitochondrial genes while RpoT3 transcribes chloroplast genes. The N-terminal amino acid composition of RpoT2 permits entry of this enzyme into both chloroplasts and mitochondria (Hedtke *et al.*, 2000; reviewed by Liere and Börner, 2007). This is possible because the RpoT2 mRNA contains two in-frame AUG codons near the 5' terminus (Richter *et al.*, 2002). Initiation of translation initiation at the second AUG codon produces a mitochondria-target translation product. In contrast, initiation of translation from the first AUG codon produces a translation product with chloroplast transit sequences.

With respect to mRNA, there are some interesting differences between monocot and dicot transcripts. For example, there are significant differences in the 5' untranslated leader sequence (located between the 5' cap and the initiation codon) with respect to GC content, length, and context of the start codon used to initiate translation. This is important because the features of the RNA that precede the initiation codon are known to affect the efficiency of translation (Pain, 1996). In general, mRNAs with longer leader sequences tend to be translated at higher rates than short-leader sequence mRNAs (Kozak, 1991; Futterer and Hohn, 1996), though this is not a hard-and-fast rule. Dicot mRNA leader sequences also tend to be AU-rich (Joshi, 1987), and many range from 10–120 nt.

In contrast, the leader sequences associated with monocot mRNAs are C-rich, ranging from about 40–120nt long. By comparison with their counterparts in the animal kingdom, plant mRNAs share many structural features and, in many cases demonstrates remarkable stability under a variety of conditions.

Finally, it is important to keep in mind that both the mitochondria and the chloroplasts in the cell are transcriptionally active compartments. Each of these organelles has its own genome, and genes are often encoded on both strands of the organellar chromosome. Proteins produced in the mitochondria remain in the mitochondria, though several additional proteins required for mitochondrial function are imported from the cytoplasm. Chloroplast genomes tend to be much larger than mtDNA and, in a manner identical to the mitochondrial proteins, chloroplast-encoded proteins remain in that organelle. Also in manner similar to mitochondria, requisite proteins are imported into the chloroplast from the cytoplasm to complete the assembly of proteins exhibiting quaternary structure. In this regard, perhaps the most noteworthy example is RUBISCO (ribulose-1,5-biphosphate carboxylase/oxygenase), which is the first of many enzymes in the photosynthetic pathway. Widely touted as the most abundant protein in the world, RUBISCO consists of eight nuclear-encoded peptides and eight chloroplast-encoded peptides and is an excellent example of cooperation among two subcellular compartments in the regulation of gene expression. For a recent review and overview of the regulation of gene expression in plants, see Farrell (2007) and Farrell and Bassett (2007).

The following sections contain examples of RNA isolation procedures that have been used successfully for a number of plant tissue types. Keep in mind that many vendors sell products that often streamline the isolation procedure and in many cases reduce or altogether eliminate toxic compounds that are traditionally associated with RNA recovery. The amount of RNA that is to be expected is highly dependent upon the plant species, the specific tissue, and the isolation procedure. For example, one gram of leaf tissue may yield 25–100 μ g total RNA while three to four grams of fruit tissue may be needed to isolate an equivalent amount of RNA.

Protocol: RNA isolation from leaf

The following protocol is a modification of the procedures of Jaakola *et al.* (2001) and Chang *et al.* (1993). It is very effective for tissues and organs with high RNAase- and phenolic compound content.

1. Flash-freeze leaf tissue in liquid nitrogen and, if not ready for immediate use, store at -80° . Leaf tissue harvested in the field should ideally be flash-frozen on site. If this is not possible, the leaf tissue should be transported back to the lab on ice and then processed immediately.
2. Grind the frozen tissue with a mortar and pestle which have been pre-cooled with liquid nitrogen. Be sure to work quickly so that neither the mortar and pestle nor the tissue has an opportunity to warm.

Note 1: After grinding, leaf tissue can be stored at -80° until ready for use.

Note 2: If more liquid nitrogen is required during grinding to avoid any thawing of the sample, be sure to add it very slowly to the mortar to avoid potentially dangerous splashing and sample loss. Wear gloves and protective eyewear at all times.

3. Weigh pulverized frozen tissue in a weight boat that was pre-cooled on dry ice to ensure a suitable mass of starting material. Transfer the frozen tissue to a pre-chilled tube on ice.
4. For each 100 mg pulverized tissue, add 750 μ l extraction buffer (2% CTAB; 2% PVP-40²; 10 mM Tris, pH 8.3; 10 mM EDTA; 2.0 M NaCl; 0.05% spermidine; add β -ME to 100 mM just prior to use) that has been preheated to 65° . Incubate tubes at 65° for 10 min.

Note 1: This procedure can be performed in a microfuge tube or scaled up, as needed. As always, be sure to check for buffer and solvent compatibility with all centrifugation tubes.

Note 2: In order to maximize the yield of RNA, vortex the tube 3–4 times during the incubation period.

5. Centrifuge samples at $10,000 \times g$ for 10 min. If possible, perform this centrifugation at 4° .
 6. Transfer the supernatant to a fresh pre-chilled tube, taking care not to disturb the pelleted debris.
 7. Add an equal volume of chloroform:isoamyl alcohol (24:1) and mix carefully.
 8. Centrifuge tubes for 2 min at $10,000 \times g$.
 9. Transfer the aqueous (upper) phase to a fresh tube.
 10. Add an equal volume of chloroform:isoamyl alcohol (24:1) and mix carefully.
 11. Centrifuge tubes for 2 min at $10,000 \times g$.
 12. Transfer the aqueous (upper) phase to a fresh tube.
 13. Add 0.25 vol 12 M LiCl to the aqueous material and mix thoroughly.
 14. Incubate at 4° overnight to precipitate the RNA.
 15. Collect the RNA by centrifugation at $12,000 \times g$ for 20 min. Decant the ethanol carefully.
 16. Wash the RNA pellet twice with 500 μ l aliquots of 70% ethanol. Decant each ethanol wash.
- Note: It is not necessary to centrifuge the tubes between washes unless the RNA dislodges from the side of the tube.*
17. Wash the RNA once with 1000 μ l 95–100% ethanol. Decant the ethanol carefully and allow the tube to air-dry (less than 5 min).
 18. Dissolve the RNA in 500 μ l of a solution of 1.0 M NaCl; 0.25% SDS; 10 mM Tris, pH 8.3, and 5 mM EDTA.
 19. Add an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). Mix carefully and thoroughly and then incubate for 5 min at 55° .
 20. Carefully mix again and then centrifuge for 5 min at $12,000 \times g$.
 21. Transfer the aqueous (upper) phase to a fresh tube, taking care not to disturb the protein interface, if any, or the lower organic phase.
 22. Add an equal volume of chloroform:isoamyl alcohol (24:1) to the aqueous material. Mix carefully.
 23. Centrifuge to 2 min at $12,000 \times g$.

²PVP = Polyvinylpyrrolidone; 40,000 MW.

24. Transfer the aqueous (upper) phase to a fresh tube. Add 2.2 volumes of 95–100% ethanol. Store at -80° for at least 2 h or at -20° overnight.
25. Collect precipitated RNA by centrifugation at $12,000 \times g$ for 15 min. If possible, perform this centrifugation step at 4° .
26. Decant the ethanol carefully. Wash the RNA pellet twice with 500 μ l aliquots of 70% ethanol. Decant each ethanol wash.
Note: It is not necessary to centrifuge the tubes between washes unless the RNA dislodges from the side of the tube.
27. Wash the RNA once with 1000 μ l 95–100% ethanol. Decant the ethanol carefully and allow the tube to air-dry (less than 5 min).
28. Dissolve the RNA in the smallest possible volume of nuclease-free water, and assess purity, sample integrity, and concentration as described in Chapter 6. If the RNA is not to be used immediately, store it in suitable aliquots at -80° .

Protocol: RNA isolation from bark

The following protocol is a modification of the procedures of Artlip *et al.* (1997) and Callahan *et al.* (1993). It has been used successfully by this lab group for a number of applications, including suppression subtractive hybridization (Chapter 22) for global analysis of gene expression in peach bark (Bassett *et al.*, 2006). Bark is a rather attractive tissue for RNA isolation, particularly for seasonal studies, because it is readily available in large quantities and does not separate from the plant in the manner of deciduous leaves, flowers, and fruit (abscission).

1. Collect bark tissue by gently scraping small pieces from twigs using a nuclease-free scalpel or other implement.

Note: If collected in the field, place bark samples on ice for transport back to the lab.

2. Flash-freeze the bark tissue in liquid nitrogen and then grind the frozen tissue. Be sure to pre-cool the mortar or pestle with liquid nitrogen, too, and work quickly so that neither the mortar nor pestle have an opportunity to warm.

Note 1: After grinding, bark tissue can be stored at -80° until ready for use.

Note 2: If more liquid nitrogen is required during grinding to avoid any thawing of the sample, be sure to add it very slowly to the mortar to avoid potentially dangerous splashing and sample loss. Wear gloves and protective eyewear at all times.

3. Rapidly transfer the pulverized, frozen tissue to a weight boat that was pre-cooled on dry ice to ensure a suitable mass of starting material.
4. Resuspend 1 g tissue in 20 ml of pre-heated (65°) extraction buffer (0.1 M Tris-Cl, pH 9.0; 0.1 M NaCl, 0.1% SDS, 1% PVP-40³. Just prior to use, bring to 10 mM ascorbic acid; 100 mM β -ME; 100 μ g/ml proteinase K). Carefully resuspend all particulate matter and then incubate for 10 min at 65° with gentle agitation.
5. Centrifuge samples for 10 min at $15,000 \times g$ to pellet cellular debris.
6. Carefully transfer the supernatant into a fresh tube. Add an equal volume of Tris-saturated phenol that was preheated to 65° . Mix carefully by vortexing.

³PVP = Polyvinylpyrrolidone; 40,000 MW.

Note: As always, be sure to check for buffer and solvent compatibility with all centrifugation tubes.

7. Centrifuge at $12,000 \times g$ for 10 min at room temperature.
8. Note the appearance of the tube at the end of the centrifugation. If the protein interface is abundant, then repeat steps 6 and 7. Otherwise, transfer the aqueous (upper) phase to a fresh tube and then add an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). Mix carefully by vortexing.
9. Centrifuge at $12,000 \times g$ for 10 min at room temperature.
10. Transfer the upper aqueous phase to a fresh tube and then add an equal volume of chloroform:isoamyl alcohol (24:1). Mix carefully by vortexing.
11. Centrifuge at $12,000 \times g$ for 3 min at room temperature.
12. Transfer the upper aqueous phase to a fresh tube that has been pre-chilled on ice.
13. Add 0.1 vol 3 M NaOAc, pH 5.2. Mix by gently vortexing the tube.
14. Add 0.1 vol 10% SDS and 0.1 vol 5 M NaCl. Mix by gently vortexing the tube.

Note: The addition of sodium acetate, SDS, and sodium chloride in steps 12 and 13 will drive the precipitation of polysaccharides when incubated on ice in the next step.

15. Incubate tubes on ice for 2 h.
16. Centrifuge at $15,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .
17. Carefully transfer the supernatant to a fresh tube. Add 0.2 vol 10 M LiCl, and mix by gently vortexing the tube.

Note: Be careful to avoid disturbing the pelleted material.

18. Incubate at -20° overnight to precipitate the RNA.
19. Centrifuge at $15,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .
20. Carefully remove the supernatant as completely as possible, taking care not to disturb the pellet.
21. Resuspend the pellet in a minimum volume of nuclease-free water. Aggressive vortexing may be needed to dissolve the pellet completely.

Note: If possible perform this step and all remaining steps in microcentrifuge tubes.

22. Add 0.02 vol 5 M NaCl and 2.5 vol ice-cold 95% ethanol. Mix thoroughly by gentle vortexing and then incubate at -20° overnight or for 3 h at -80° .
23. Centrifuge at $12,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .
24. Remove the ethanol supernatant as completely as possible. Wash the pellet twice with 1 ml 70% ethanol to remove the excess salt.

Note: It is not necessary to centrifuge the sample between washes unless the precipitate dislodges from the side of the tube.

25. Wash the precipitate once with 1 ml 100% ethanol, and then remove it as completely as possible.
26. Allow the sample to air-dry (less than 5 min), but do not dry it out completely.

Note: Keeping the RNA slightly damp with ethanol will facilitate dissolving it aqueous buffer.

27. Resuspend the RNA in a small volume of nuclease-free water and allow the sample to dissolve on ice. Pipette up-and-down thoroughly to create a homogeneous solution.
28. Measure concentration and estimate purity spectrophotometrically, as described in Chapter 6. Store RNA in suitable aliquots at -80° until ready to use.

Protocol: RNA isolation from fruit

The following protocol for the isolation of RNA from fruit is a modification of the procedures of Callahan *et al.* (1989) and Morgens *et al.* (1990). Plant

tissue is frozen in liquid nitrogen immediately upon harvesting and then stored at -80° until ready to use. Alternatively, fruit tissue can be lyophilized (may require days) prior to commencing the isolation protocol, though RNA quality may be compromised (A. Callahan, personal communication).

1. Grind frozen tissue in liquid nitrogen using a pre-cooled mortar and pestle. Do not allow the mortar, pestle or tissue to warm.

Note: If more liquid nitrogen is required during grinding to avoid any thawing of the sample, be sure to add it very slowly to the mortar to avoid potentially dangerous splashing and sample loss. Wear gloves and protective eyewear at all times.

2. Rapidly transfer the pulverized tissue to a weight boat that was pre-cooled on dry ice to ensure a suitable mass of starting material.
3. In a 50-ml Oak Ridge tube, resuspend up to 1 g tissue in 20 ml of extraction buffer (0.1 M Tris-Cl, pH 9.0; 0.1 M NaCl, 0.1% SDS, 1% PVP-40⁴ that was preheated to 65° . Just prior to use, bring to 10 mM ascorbic acid; 100 mM β -ME; 100 μ g/ml proteinase K). Carefully resuspend all particulate matter.

Note: The mass of starting material is variable depending on the tissue type. For example, use 0.3 g if working with leaf tissue, 0.5 g if working with early fruit, or 2 g if working with very ripe fruit.

4. Centrifuge for 10 min at $12,000 \times g$.

Note: Be attentive to the manufacturer's recommendations pertaining to the proper use of Oak Ridge tubes.

5. Carefully transfer the supernatant into a fresh Oak Ridge tube contained 20 ml of Tris-saturated phenol that was preheated to 65° . Mix carefully by vortexing.

Note: As always, be sure to check for buffer and solvent compatibility will all centrifugation tubes.

6. Centrifuge at $12,000 \times g$ for 10 min at room temperature.
7. Examine the contents of the Oak Ridge tube. If the protein interface is abundant, then repeat steps 5 and 6. Otherwise, transfer the upper aqueous material to a fresh Oak Ridge tube containing 18 ml phenol:chloroform:isoamyl alcohol (25:24:1). Mix carefully by vortexing.
8. Centrifuge at $12,000 \times g$ for 10 min at room temperature.
9. Transfer the (upper) aqueous material to a fresh Oak Ridge tube containing 18 ml chloroform:isoamyl alcohol (24:1) using a nuclease-free pipette. Mix carefully by vortexing.
10. Centrifuge at $12,000 \times g$ for 3 min at room temperature.
11. Transfer the (upper) aqueous phase to a fresh Oak Ridge tube that has been pre-chilled on ice.
12. Add 0.1 vol 3 M NaOAc, pH 5.2. Mix by gently vortexing the tube.
13. Add 0.05 vol 20% SDS and 0.1 vol 5 M NaCl. Mix by gently vortexing the tube.
14. Incubate tubes on ice for 2 h.
15. Centrifuge at $12,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .
16. Carefully transfer the supernatant to a fresh tube. Add 0.2 vol 10M LiCl, and mix by gently vortexing the tube.
17. Incubate at -20° overnight to precipitate the RNA.
18. Centrifuge at $12,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .
19. Carefully remove the supernatant as completely as possible, taking care not to disturb the pellet.

⁴PVP = Polyvinylpyrrolidone; 40,000 MW.

20. Resuspend the pellet in a minimum volume of nuclease-free water, keeping in mind that more than 1 ml may be required. Aggressive vortexing may be needed to dissolve the pellet completely.

Note: If possible perform this step and all remaining steps in microcentrifuge tubes.

21. Add 0.02 vol 5 M NaCl and 2.5 vol ice-cold 95% ethanol. Mix thoroughly by gentle vortexing and then incubate at -20° overnight.

22. Centrifuge at $12,000 \times g$ for 20 min. If possible, perform this centrifugation at 4° .

23. Remove the ethanol supernatant as completely as possible. Wash the pellet twice with 1 ml 70% ethanol to remove the excess salt.

Note: It is not necessary to centrifuge the sample between washes unless the precipitate dislodges from the side of the tube.

24. Wash the precipitate once with 1 ml 100% ethanol, and then remove it as completely as possible.

25. Allow the sample to air-dry, but do not dry it out completely.

Note: Keeping the RNA slightly damp with ethanol will facilitate dissolving it aqueous buffer.

26. Resuspend the RNA in 100–500 μ l nuclease-free water and allow the sample to dissolve on ice. Pipette up-and-down thoroughly to create a homogeneous solution.

27. Measure concentration and estimate purity spectrophotometrically, as described in Chapter 6. Store RNA in suitable aliquots at -80° until ready to use.

Strategies for RNA isolation from other plant tissues

One of the primary complaints associated with RNA isolation from plant tissue is an unacceptably low yield. There are several reasons why this could happen. First, and most assuredly, the plant cell wall is always a challenge. While it is true that the wall can be dismantled enzymatically, the process is rather inefficient and time-consuming. It is also true that a fair amount of damage can be inflicted upon plant tissue by dropping it into a chaotropic lysis buffer and homogenizing it, though a large number of cells may remain intact, reducing the yield of RNA that might otherwise be possible. Moreover, some tissues are rather reluctant to give up their RNA. Thus, a majority of procedures in this domain begin with pulverizing the starting tissue in liquid nitrogen down to a fine powder. When in doubt, pulverize. In instances where it is not practical to maintain field-harvested tissue in a frozen state until returning to the lab, preparing a guanidinium lysate on-site is usually a satisfactory alternative.

Second, tissues from some plants, including pine, grape, and tea leaves contain high concentrations of phenolics⁵. Elimination of these compounds is essential because phenolic compounds form quinones upon oxidation (Loomis, 1974) which crosslink nucleic acids, rendering them useless for molecular biology applications. Moreover, endogenous RNase is also a common problem. It is also possible that RNase may be accidentally introduced by the investigator.

⁵Polyphenols tend to be more prevalent in conifers than in deciduous trees (Schneiderbauer *et al.*, 1991).

A surprising source of RNase contamination in the isolation of plant RNA are the mortar and pestle, both of which should be squeaky-clean when used to grind flash-frozen plant tissue. One easy way to ensure RNase neutralization is to soak the mortar and pestle in H_2O_2 for 15–20 min, as described in Chapter 7, followed by several rinses with nuclease-free water. It is possible to eliminate toxic compounds such as phenol and guanidinium buffers altogether and still be able to rapidly recover total RNA (Chang *et al.*, 1993) using a CTAB lysis system, though chloroform is used to remove the proteins.

Third, high concentrations of polysaccharides, waxes, and sometimes latex, can be problematic. Occasionally, plant RNA isolation protocols raise eyebrows among investigators who previous have not worked with plant tissue. Certain protocols incorporate three or more precipitations of the RNA, particularly Tris-lithium chloride procedures, as a means of eliminating the massive amount of carbohydrate that fruit, in particular, can render. Some investigators actually rinse pelleted RNA with 3 M NaOAc because of the ability of this solution to wash away polysaccharide contaminants (Logemann *et al.*, 1987; Lal, 2001). It is true that minimizing the number of manipulations to which the RNA is subject is quite favorable, but it is also true that contaminants which copurify with the RNA must be dealt with assertively. Plant RNA lysis buffers also often feature the inclusion of spermidine (as high as 0.5% w/v) and/or PVP (up to 2% w/v). Both of these compounds can be quite helpful for maximizing high purity RNA recovery. While well-known for its ability to enhance RNA polymerase activity, spermidine in this case forms a low-solubility complex with genomic DNA in the lysate, making it easier to remove the DNA. PVP is a rather large polysaccharide that essentially takes up physical space, thereby increasing the effective concentration of other molecules and making it easier to precipitate them. Further, PVP and β -ME, another common lysis buffer component, are effective reducing reagents, which will minimize quinone formation due to the oxidation of phenolic compounds.

The good news is that plant RNA in general is quite a bit more stable than RNA isolated from animal cells and, according to some investigators, is legendary in terms of capacity for long-term storage. For example, we have cloned genes and mapped transcription start sites using morning glory RNA that had been stored at -80° for years.

Troubleshooting RNA isolation from plant tissue

1. When working with plant RNA, low A_{260}/A_{230} may indicate polysaccharide contamination. Consequently, the RNA isolation protocol may need to be modified to ensure polysaccharide elimination. Further, a low A_{260}/A_{280} ratio is an indicator of protein contamination. See Chapter 6 for a detailed discussion.
2. As certain plant tissues mature, particularly leaf, the yield of RNA tends to decrease, while certain plants, e.g. *Arabidopsis*, tend to render less RNA in general, compared to other species.

3. If the RNA yield is low, it could also be due to inefficient homogenization or incomplete dispersal of pulverized tissue. If the former is suspected, one should examine a few microliters of the lysate under a microscope to assess the level of cell disruption and then adjust the procedure accordingly. If the latter is suspected, be sure that all pulverized tissue is dispersed uniformly before continuing on with the isolation procedure.
4. In the instance of plant mRNA isolation, multiple precipitations with LiCl are particularly valuable for the selective precipitation of RNA when polysaccharide contaminants are present.
5. A very nice method for removing impurities and concentrating a sample, particularly if using homemade reagents, is to use one of the very popular silica-based spin columns. Briefly, dissolve the RNA in the chaotropic (guanidinium) lysis buffer that accompanies the kit, bind the RNA to column, wash the column, and then elute in nuclease-free water, according to the manufacturer's instructions.
6. Degraded RNA could be a sign that (a) the tissue had high levels of endogenous RNase that was not controlled; (b) the sample thawed prior to addition of the lysis/extraction solution; or (c) RNase was accidentally introduced by the user, meaning that a modification in technique is in order.
7. As indicated above, the best strategy for RNA isolation from a particular plant tissue will need to be determined empirically. Kits for plant RNA isolation abound, and one should not be shy about asking vendors for a first-time evaluation (free) kit.

References and suggested reading

- Ainsworth, C. (1994). Isolation of RNA from flora tissue of *Rumex acetosa*. *Plant Mol. Biol. Rep.* **12**, 198–203.
- Artlip, T.S., A.M. Callahan, C.L. Bassett and M.E. Wisniewski. (1997). Seasonal expression of a dehydrin gene in sibling deciduous and evergreen genotypes of peach (*Prunus persica* [L]). Batsch. *Plant Mol. Biol.* **33**, 61–70.
- Bassett, C.L., M.E. Wisniewski, T.S. Artlip, G.A. Richart, J.L. Norelli, and R.E. Farrell, Jr. (2009). Comparative expression and transcript initiation of three peach dehydrin genes. *Planta* **230**, 107–118.
- Bassett, C.L., M.E. Wisniewski, T.S. Artlip, J.L. Norelli, J. Renaut and R.E. Farrell, Jr. (2006). Global analysis of genes regulated by low temperature and photoperiod in peach bark. *J. Am. Soc. Hort. Sci.* **131**, 551–563.
- Bassett, C.L., M.L. Nickerson, R.E. Farrell, Jr., T.S. Artlip, A.E. Ghaouth, C.L. Wilson and M.E. Wisniewski. (2005). Characterization of an S-locus receptor protein kinase-like gene from peach. *Tree Physiol.* **25**, 403–411.
- Bassett, C.L., M.L. Nickerson, R.E. Farrell, Jr. and M. Harrison. (2004). Multiple transcripts of a leucine-rich repeat receptor kinase from morning glory originate from different TATA boxes in a tissue-specific manner. *Mol. Gen. Genom.* **271**(6), 752–760.
- Bilgin, D.D., E.H. DeLucia and S.J. Clough. (2009). A robust RNA isolation method suitable for Affymetrix GeneChip analysis and quantitative real-time PCR. *Nature Protocols* **4**, 333–340.
- Callahan, A., P. Morgens and E. Walton. (1989). Isolation and *in vitro* translation of RNAs from developing peach fruit (*Prunus persica* L. Batsch 'Reliance'). *HortScience* **24**, 356–358.

- Callahan, A., P. Morgens and R. Cohen. (1993). Isolation and initial characterization of cDNAs for mRNAs regulated during peach fruit development. *J. Am. Hort. Soc.* **118**, 531–537.
- Chang, S., J. Puryear and J. Cairney. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* **11**, 113–116.
- Farrell, Jr., R.E. (2007). The regulation of gene expression in plants and animals. In *Regulation of Gene Expression in Plants: the Role of Transcript Structure and Processing* (C.L. Bassett Ed.). Springer Publishers. pp. 1–38.
- Farrell, Jr., R.E. and C.L. Bassett. (2007). Multiple transcript initiation as a mechanism for regulating gene expression. In *Regulation of Gene Expression in Plants: The Role of Transcript Structure and Processing* (C.L. Bassett Ed.). Springer Publishers. pp. 39–66.
- Futterer, J. and T. Hohn. (1996). Translation in plants – rules and exceptions. *Plant Mol. Biol.* **32**, 159–189.
- Graham, G.C. (1993). A method for extractions of total RNA from *Pinus radiata* and other conifers. *Plant Mol. Biol. Rep.* **11**, 32–37.
- Green, M.J., D.A. Thompson and D.J. MacKenzie. (1999). Easy and efficient DNA extraction from wood plants for the detection of phytoplasmas by polymerase chain reaction. *Plant Dis.* **83**, 482–485.
- Hedtke, B., T. Börner and A. Weihe. (2000). One RNA polymerase serving two genomes. *EMBO Reports* **51**, 435–440.
- Jaakola, L., A.M. Pirttilä, M. Halonen and A. Hohtola. (2001). Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit. *Molecular Biotechnology* **19**, 201–203.
- Jaiprakash, M.R., B. Pillai, P. Venkatesh, N. Subramanian, V.P. Sinkar and P.P. Sadhale. (2003). RNA isolation from high-phenolic freeze-dried tea (*Camellia sinensis*) leaves. *Plant Mol. Biol. Rep.* **21** 465a–465g.
- Joshi, C.P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.* **15**, 6643–6652.
- Kanazawa, A., M. O'Dell, R.P. Hellens, E. Hitchin and M. Metzloff. (2000). Mini-scale method for nuclear run-on transcription assay in plants. *Plant Mol. Biol. Rep.* **18**, 377–383.
- Kanno, T., B. Huettel, M.F. Mette, W. Aufsatz, E. Jaligot, L. Daxinger, D.P. Kreil, M. Matzkel and A.J.M. Matzke. (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nature Genet.* **37**, 761–765.
- Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867–19870.
- Lal, L., R. Sahoo, R. Kumar-Gupta, P. Sharma and S. Kumar. (2001). RNA isolation from high-phenolic tea leaves and apical buds. *Plant Mol. Biol. Rep.* **19** 181a–181f.
- Lewinsohn, E., C.L. Steele and R. Croteau. (1994). Simple isolation of functional RNA from woody stems of gymnosperms. *Plant Mol. Biol. Rep.* **12**, 20–25.
- Liere, K. and T. Börner. (2007). Transcription and transcriptional regulation in plastids. In *Topics in Current Genetics: Cell and Molecular Biology of Plastids* (R. Bock Ed.). Springer, New York. pp. 121–174.
- Logemann, J., J. Schell and L. Willmitzer. (1987). Improved method for the isolation of RNA from plant tissues. *Analyt. Biochem.* **163**, 16–20.
- Loomis, W. (1974). Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. *Methods Enzymol.* **31**, 528–545.

- López-Gómez R. and M.A. Gómez-Lim. (1992). A method for extracting intact RNA from fruits rich in polysaccharides using ripe mango mesocarp. *Hort. Sci.* **27**, 440–442.
- Mitra, D. and A. Kootstra. (1993). Isolation of RNA from apple skin. *Plant Mol. Biol. Rep.* **11**, 326–332.
- Morgens, P., A. Callahan, L.J. Dunn and F. Abeles. (1990). Isolation and sequencing of a cDNA clone encoding an ethylene-induced peroxidase from cucumber cotyledons. *Plant Mol. Biol.* **14**, 715–725.
- Onodera, Y., J.R. Haag, T. Ream, P.C. Nunes, O. Pontes and C.S. Pikaard. (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622.
- Norelli, J.L., R.E. Farrell, Jr., C.L. Bassett, A.M. Baldo, D.A. Lalli, H.S. Aldwinckle and M.E. Wisniewski. (2009). Rapid transcriptional response of apple to fire blight disease revealed by cDNA suppression subtractive hybridization analysis. *Tree Genet. Genomes* **5**, 27–40.
- Pain, V. (1996). Initiation of protein synthesis in eukaryotic cells. *Eur. J. Biochem.* **236**, 747–771.
- Palmiter, R.D. (1974). Magnesium precipitation of ribonucleoprotein complexes: expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* **13**, 3606.
- Richter, U., J. Kiessling, B. Hedtke, E. Decker, R. Reski, T. Borner and A. Weihe. (2002). Two RpoT genes of *Physcomitrella patens* encode phage-like RNA polymerases with dual targeting to mitochondria and plastids. *Gene* **290**, 95–105.
- Schneiderbauer, A., H. Sandermann, Jr. and D. Ernst. (1991). Isolation of functional RNA from plant tissues rich in phenolic compounds. *Analyt. Biochem.* **197**, 91–95.
- Soni, R. and J.A.H. Murray. (1994). Isolation of intact DNA and RNA from plant tissues. *Analyt. Biochem.* **218**, 474–476.
- Strommer, J., R. Gregerson and M. Vayada. (1993). Isolation and characterization of plant mRNA. In *Methods in Plant Molecular Biology and Biotechnology* (B.R. Glick and J.E. Thompson, Eds). CRC Press, Boca Raton, FL.
- Wan, C.Y. and T.A. Wilkens. (1994). A modified hot borate method significantly enhances the yield of high quality RNA from cotton. *Analyt. Biochem.* **223**, 7–12.
- Wisniewski, M.E., C.L. Bassett, J. Renaut, R.E. Farrell, Jr., T. Tworcoski and T.S. Artlip. (2006). Differential regulation of two dehydrin genes from peach (*Prunus persica* L. Batsch) by photoperiod, low temperature, and water deficit. *Tree Physiol.* **26**, 575–584.

This page intentionally left blank

5 Isolation of Polyadenylated RNA

Rationale	121	Oligo(dT)-cellulose column chromatography	129
Polyadenylation	122	Protocol: purification of biophysical quantities of poly(A) ⁺ RNA	130
The poly(A) caveat	123	Rapid, non-column poly(A) ⁺ purification	135
Example 1	124	Protocol: non-column poly(A) ⁺ purification	135
Example 2	124	References	136
Selection of polyadenylated molecules	126		
Magnetic bead technology for poly(A) ⁺ purification	127		

Rationale

Why purify mRNA? Recall that only a small fraction of all mature transcripts in the cytoplasm are mRNA, usually on the order of two to three percent. The underlying rationale for purifying mRNA, i.e. remove the rRNA and tRNA species, is to increase the statistical representation of one, several, or all of the mRNAs in a sample as a function of the total mass of purified RNA. Any approach of this nature is known as an enrichment strategy. One form of enrichment includes the manipulation of biological material prior to cellular disruption to superinduce, or accumulate, certain types of transcripts in the cell. More commonly, enrichment involves the manipulation of a purified RNA sample to eliminate as much of the non-mRNA as possible. Reasons for enriching samples in favor of mRNA may include, but are certainly not limited to: (1) increasing the sensitivity of transcription assays such as Northern analysis, nuclease protection, or RT-PCR; or (2) producing a fully representative cDNA¹. By increasing statistical representation one increases the likelihood of being able to identify rare mRNA sequences on a blot, in a library, or by RT-PCR.

It is important to understand, however, that given the refinements in several reagents and techniques, it may no longer be necessary or even desirable to work with the mRNA fraction. Because of improvements in assay sensitivity and the requirements for remarkably small amounts of RNA (often nanogram

¹cDNA synthesis may be performed for traditional library synthesis, transcript measurement by PCR, microarray probe synthesis, or other related purposes.

quantities), mRNA selection may further under-represent low abundance transcripts (discussed below) and compromise the accuracy of the data. In this laboratory, all of the transcription assays performed, including quantitative and semi-quantitative PCR, begin with total cellular or total cytoplasmic RNA as template material.

Eukaryotic mRNAs are categorized as falling into an abundance class based on prevalence (i.e., the average number of copies of an RNA transcript) in the cell. High abundance mRNAs² are those that are present in hundreds of copies per cell. If a cell produces large amounts of a particular protein, then one might reasonably expect that the cell transcribes a correspondingly elevated mass of the message. For example, reticulocytes and oviduct cells were selected to clone β -globin (Efstratiadis *et al.*, 1976) and ovalbumin (Buell *et al.*, 1978), respectively. Medium abundance mRNAs are those present in dozens of copies per cell. Many of the standard housekeeping genes (e.g., histone, β -actin, GAPDH), the expression of which are assayed frequently as control or reference genes in RNA-based analyses, produce transcripts of sufficient number to be in this class. Low abundance mRNAs are those transcripts that are present in 14 or fewer copies per cell (Williams, 1981); these are the so-called “rare” mRNAs, the detection of which has been enhanced greatly through the application of such super-sensitive assays as the polymerase chain reaction (PCR). It has been estimated that fully 30% of cellular mRNAs fall in the low abundance category. Further, this author’s definition of *very low abundance* transcripts are those present, on average, in fewer than one copy per cell. In the old days of molecular biology transcripts such as these were referred to, affectionately, as the “hard to clone” genes!

Polyadenylation

The existence of the poly(A) tail has been known for several years. This structure is a tract of adenosine nucleotides that characterizes the 3' end of most, but not all, eukaryotic mRNA molecules. Polyadenylation, the process by which this structure is added to mRNA, is not coupled directly to transcription termination (Darnell, 1979; Manley, 1988); rather it is catalyzed by a family of nuclear enzymes (Moore and Sharp, 1985), known collectively as poly(A) polymerase. Thus, polyadenylation *in vivo* is not an integral part of transcription, but rather a very rapid (Nevins and Darnell, 1978) posttranscriptional regulatory event.

The length of the poly(A) tail is variable among mRNAs, even among those RNA molecules that were synthesized by transcription of the same locus. The typical mammalian poly(A) tail averages 200 nucleotides (nts) long, though these tracts may range from 50–300 nts. All RNA molecules that possess a poly(A) tail are referred to collectively as poly(A)⁺ mRNA and are correctly said to be polyadenylated.

²Certain non-mRNA transcripts are also classified as being highly abundant in the cell, specifically rRNA and tRNA.

At least one function of the poly(A) tail is regulating the stability and functionality of the mRNA (reviewed by Ross, 1995), and length of the poly(A) tail is known to shorten as it ages. *In vivo*, the poly(A) tract associates with poly(A)-binding protein (PABP), which binds stoichiometrically to the poly(A) tail every 15–20 bases (reviewed by Kramer, 1996). This interaction of the poly(A) tail with PABP helps the transcript resist 3' → 5' nuclease attack. Of course, PABP and other proteins dissociate from the RNA during the course of most RNA isolation procedures. The poly(A) tail also has a role facilitating transport of the mature mRNA out of the nucleus (nucleocytoplasmic transport), and although the details are unclear, it is known that the adenosine analog 3'-deoxyadenosine (cordycepin) inhibits polyadenylation of nuclear RNA and curtails the appearance of these transcribed, nonadenylated sequences in the cytoplasm.

The generation of the 3' poly(A) tract is actually the product of two sequential enzymatic activities, namely a cleavage reaction of the precursor RNA (hnRNA) followed by the actual polyadenylation event; primary RNA transcripts may extend up to thousands of bases beyond the polyadenylation signal (Nevins and Darnell, 1978; Ziff, 1980). The precise site at which the poly(A) tail is added is formed by an endonucleolytic cleavage of the precursor RNA molecule (Nevins and Darnell, 1978), and not by exonucleolytic degradation from the 3' end of the molecule generated upon transcription termination. This cleavage event is then followed by the sequential addition of the adenylate residues that ultimately constitute the poly(A) tail.

The cleavage that precedes polyadenylation is not a haphazard occurrence; the efficiency and location of this event are precisely regulated by the aptly named polyadenylation signal. Among the higher eukaryotes, virtually all polyadenylated mRNA molecules exhibit the highly conserved hexanucleotide motif 5'...AAUAAA...3', or a closely related variant, which can typically be identified fewer than 30 bases upstream from the polyadenylation start site (Proudfoot and Brownlee, 1976). Efficient polyadenylation also requires the presence of a second element, a downstream GU-rich cluster (Gil and Proudfoot, 1984; McDevitt *et al.*, 1984; Sadofsky and Alwine, 1984; Cole and Stacy, 1985; Conway and Wickens, 1985). This GU-rich sequence is less highly conserved than the AAUAAA motif and is found approximately 20–40 nucleotides downstream from the actual cleavage site. Polyadenylation subsequently proceeds via the formation of a heteromultimeric complex (Takagaki *et al.*, 1988) consisting of endonuclease components, a GU-rich binding factor, and poly(A) polymerase itself.

The poly(A) caveat

Perhaps the most important point of this entire chapter pertains to the significance of mRNA that is purified by some type of affinity chromatographic selection directed toward the poly(A) tail. Although most eukaryotic mRNAs are

polyadenylated, a small percentage is not. Thus, populations of poly(A)-selected mRNAs molecules do not include non-polyadenylated mRNAs, an element that may or may not be meaningful in the context of a particular investigation.

Example 1

Northern analysis of poly(A)⁺-selected mRNA could not be utilized to derive qualitative or quantitative information about cellular or cytoplasmic poly(A)⁻ species, such as the histone mRNAs. The situation becomes more complicated when one considers the possibility that alternative posttranscriptional processing pathways might yield a collection of polyadenylated and non-polyadenylated transcripts from the same locus.

The reader is cautioned that kits that purport to facilitate the selection of “total mRNA” through poly(A) selection will fail to yield poly(A)⁻ mRNA. Further, because the poly(A) tail is added in the nucleus, selection of poly(A)⁺ RNA from a whole cell lysate will unquestionably render polyadenylated cytoplasmic RNA (mRNA) as well as polyadenylated nuclear RNA (hnRNA); the mixture of the two cannot be construed as representing functional (translatable) mRNA. Recall that a great deal of hnRNA never matures into cytoplasmic mRNA and is eventually degraded in the nucleus. In order to focus more closely on the mass of poly(A)⁺ mRNA being produced by the cell, one could lyse the cells gently with hypotonic, NP-40 buffer (Chapter 2), remove the intact nuclei, and then proceed to select the polyadenylated RNA from the cytosol. Alternatively, to give even clearer definition to those specific mRNAs destined to be translated, one could prepare the polysome fraction to determine which RNAs are actually engaged by the ribosomes (Chapter 3). In one version of this approach, a technique known as the EDTA-release assay, the chelation of Mg²⁺ by EDTA causes dissociation of ribosomes into their individual subunits and release of the RNA that was undergoing translation at that very moment (Aziz and Munro, 1986; Meyuhas *et al.*, 1987; reviewed by Pierandrei-Amaldi and Amaldi, 1994). These, and more recent methods (Meyuhas *et al.*, 1996), are very useful for discerning the active mRNA population.

Example 2

A common approach to the synthesis of cDNA for RT-PCR analysis or the synthesis of cDNA libraries begins with the selection of poly(A)⁺ mRNA. To initiate conversion of mRNA into first-strand cDNA, an oligo(dT) primer is annealed to the mRNA template to provide the requisite 3'-OH group in the proper orientation. The net effect is that reverse transcriptase³ initiates first-strand cDNA synthesis at or near the 3' end of the mRNA. Scientists should

³RNA-dependent DNA polymerase. The role of reverse transcriptase is discussed in depth in Chapter 17.

be aware that cDNA libraries prepared in this fashion do not contain clones corresponding to poly(A)⁻ mRNA. In order to circumvent the potential shortcomings of oligo(dT)-primed cDNA populations, newer libraries and more inclusive cDNA pools are prepared by random-primed synthesis of the first strand cDNA, and sometimes generated through the use of random primers *and* oligo(dT) in the same reaction tube, which is a standard procedure in this laboratory. Briefly, a number of nucleotide hexamers or nonomers of random nucleotide sequence are allowed to anneal to heat-denatured RNA. This annealing effectively produces more than one 3'-OH primer upon which reverse transcriptase can act simultaneously. This will also occur regardless of whether the 3' terminal structure of the molecule is poly(A)⁺ or poly(A)⁻. cDNA synthesized by random priming alone is unlikely to include sequence corresponding to the poly(A) tail unless an oligo(dT) random primer had been included in the primer mixture. Moreover, random priming generally produces a greater number of full length cDNA, thereby facilitating more accurate analysis by any of a variety of classical or contemporary methods. When screening a library that may have been synthesized some years ago, knowledge of the method of cDNA priming may influence interpretation of the outcome of the experiment⁴.

Succinctly, then, there are several advantages and disadvantages to selecting poly(A)⁺ RNA (Table 5.1). One must weigh the relative merits of working with

Table 5.1 Poly(A)⁺ Purification: Advantages and Disadvantages

Advantages

Poly(A)⁺ purification increases mRNA mass as a percentage of all RNA in the sample.
 Poly(A)⁺ purification results in minimal interference from rRNA and tRNA.
 Poly(A)⁺ purification increases assay sensitivity.
 Magnetic methods of poly(A) capture help to concentrate the sample dramatically without the need to reprecipitate the sample.
 Purified poly(A)⁺ mRNA is ready for immediate use.
 The poly(A)⁻ fraction, which can be collected and precipitated, is an excellent negative control, especially for blot analysis.

Disadvantages

Poly(A)⁺ selection excludes poly(A)⁻ mRNAs.
 Loss of poly(A)⁺ during the purification process may further under-represent rare mRNA, especially if starting with a previously isolated sample of total RNA.
 Depending on lysis method, poly(A)⁺ purification may also yield polyadenylated hnRNA.
 Oligo(dT) matrices are expensive.
 Given the improvements in cloning and RT-PCR efficiency in the last 5 years, poly(A)⁺ selection may not be necessary.

⁴In some extreme cases, total RNA has been treated with poly(A) polymerase to polyadenylate all RNA prior to cDNA synthesis! This is neither necessary nor recommended.

poly(A)⁺ instead of total RNA in the context of the required level of sensitivity and the number of cells or mass of tissue available. In many cases, it may be worth attempting an assay using cellular RNA first, before incurring the added expense to purify poly(A)⁺ mRNA.

Selection of polyadenylated molecules

In order to separate polyadenylated from non-polyadenylated RNA species, the naturally occurring hydrogen bonding between the bases adenine and thymine can be exploited to isolate the relatively small poly(A)⁺ mRNA fraction from total cytoplasmic or total cellular RNA. Fractionation in this manner is a fine example of affinity chromatography, that is, a separation or chromatography based on a biological activity or structure. In this case, the biological structure is the poly(A) tail. This is performed with oligo(dT) linked to a variety of support matrices. For example, the various grades of oligo(dT)-cellulose consist of oligodeoxythymidylate residues (oligo(dT)) covalently bound by their 5' phosphate to cellulose beads (Gilham, 1964; Aviv and Leder, 1972). Thus, the poly(A) tail forms hydrogen bonds with the oligo(dT), thereby retaining it on the matrix. Although no longer the most commonly used matrix, oligo(dT)-celluloses remain commercially available from a number of suppliers. While oligo(dT)-cellulose has proven to be fairly effective for the resolution and detection of poly(A)⁺ mRNA, newer methods, described below, which require minimal amounts of starting material are now preferred and are in widespread use.

Although the separation mechanics have changed, poly(A) selection procedures remain variations on the theme of affinity separation. These approaches have emerged to streamline the somewhat cumbersome mechanics of what was once known as “running an oligo(dT) column” (discussed below). Alternative approaches include the following:

1. Using paramagnetic beads to which oligo(dT) has been covalently linked. Following solution hybridization, the hybrids are sequestered using a magnet.
2. Using oligo(dT) linked to latex and/or polystyrene beds in a packed column format to support chromatography.
3. Exploiting the natural affinity between biotin and avidin. Biotin can be conjugated to nucleic acids and a variety of other molecules without changing their biological activities. The tight binding of biotin by streptavidin ($K_d 10^{-15}$ M) makes many applications possible. In one format, biotinylated oligo(dT) molecules base-pair with poly(A)⁺ mRNA, and are then precipitated as a biotin/streptavidin complex. In another format, the hybrids may also be recovered using streptavidin linked to a paramagnetic bead and eluted in small, very convenient volumes without the need for further precipitation. Many permutations of this general approach are available as kits for RNA purification.
4. Mixing oligo(dT)-cellulose or, preferably, microcrystalline oligo(dT)-cellulose and total RNA together directly in a microfuge tube without any prior formation of some type of matrix geometry.

In all cases, the most important aspect of poly(A)⁺ selection is keeping the volumes involved as small as possible and, in so doing, keeping the sample as concentrated as possible.

Clearly, a variety of techniques has been developed for the purification of poly(A)⁺ mRNA and these methods have at least one thing in common: poly(A)⁺ mRNA anneals to an affinity matrix under high ionic strength (high salt) conditions. This is necessary to eliminate the electrostatic repulsion that would occur naturally between the target RNA and the thymidylate residues which are linked to the matrix. Under high-salt conditions, a monovalent cation acts as a counter-ion, neutralizing the net negative charge intrinsic to the phosphodiester backbone of RNA and DNA. In the absence of a high-salt milieu, the electrostatic repulsion would be of sufficient magnitude to prevent base-pairing and the purification of poly(A)⁺ mRNA in this manner.

The binding capacity of oligo(dT)-matrices⁵ increases with salt concentration up to about 500 mM NaCl, KCl, or LiCl, though the binding capacity is greater when KCl or LiCl is used in place of an equivalent amount of NaCl (Mercer and Naora, 1975). This increase in binding capacity, however, may be accompanied by an increase in non-specific hybridization to the matrix. Once the non-hybridized, non-poly(A)⁺ material is removed by what amounts to washing the matrix with a fairly high-salt buffer, the hybridized species of poly(A)⁺ RNA are recovered from the matrix by changing to a very low-salt buffer (0–5 mM salt). Removal of the monovalent cation from the system reestablishes electrostatic repulsion, favoring rapid dissociation of the hybridized RNA from the matrix; pre-heating the elution buffer to 55° also helps effect complete elution in a small volume. If necessary, the now separated poly(A)⁺ and poly(A)⁻ material can be concentrated by salt and alcohol precipitation. The poly(A)⁻ RNA makes an excellent negative control in all types of hybridization analyses as well as in the synthesis of cDNA (Chapters 13 and 17, respectively) because samples which have been depleted of poly(A)⁺ RNA should fail to produce a signal.

Magnetic bead technology for poly(A)⁺ purification

The covalent linkage of a variety of ligands to what are essentially iron beads has become an invaluable tool for the resolution of a variety of compounds from complex mixtures. These ligands include antibodies, which can be used to recognize and bind cell surface antigens and, in the context of this resource, oligothymidylate, for the efficient recovery of polyadenylated transcripts. Several companies now produce magnetic bead-based products; Dynabeads[®] Oligo(dT) (Invitrogen, Carlsbad, CA), for example, have been used in this lab for many years. Dynabeads come in various sizes depending on the application.

⁵ See manufacturer's package insert for the particulars of each grade and lot of oligo(dT) separation matrix.

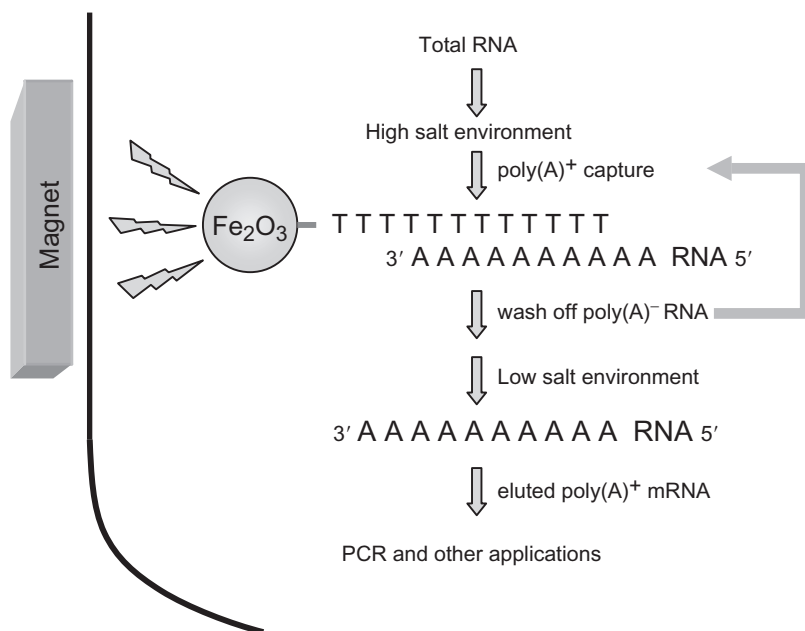


Figure 5.1 Selection of poly(A)⁺ RNA using magnetic bead technology. Under high-salt conditions previously purified RNA or a whole cell lysate is mixed in a microfuge tube with oligo(dT) linked to a paramagnetic bead. After annealing of the poly(A) tail and the oligothymidylate, polyadenylated mRNA is sequestered to the inside wall of the tube by placing in the vicinity of a magnet. After washing away nonadenylated transcripts, the purified mRNA is eluted in a very small volume of low-salt buffer or sterile water. This approach is compatible with plant and animal RNA isolated by any procedure.

The Dynabeads used for mRNA capture are uniform superparamagnetic beads, 2.8 μm in diameter, with a polystyrene coating that encloses iron oxide (Fe₂O₃) particles. These particles form a magnetic dipole when exposed to a magnetic field. Chains of deoxythymidylate, 25 nucleotides long, have been covalently attached to the surface of the beads via a 5' linker group. A mixture of total RNA added to the beads under conditions of high ionic strength will favor hybridization between polyadenylated RNA and the oligo(dT) tract linked to the beads (Fig. 5.1). Hybridization kinetics are similar to those found in free solution, with complete hybridization being observed within 5 minutes. The magnetic beads and hybridized poly(A)⁺ RNA are then collected by placing the sample tube in the vicinity of a magnet (Fig. 5.2). Washing is performed, followed by elution using a standard combination of temperature and very low-salt buffer. Isolation of polyadenylated RNA by magnetic capture precludes the requirement for traditional column affinity chromatography and has achieved a new level of resolution in the chromatography of nucleic acids. The principal advantage of this technique is the rapid isolation of pure, intact poly(A)⁺ RNA in concentrated form. As the purified poly(A)⁺ RNA is already highly concentrated,

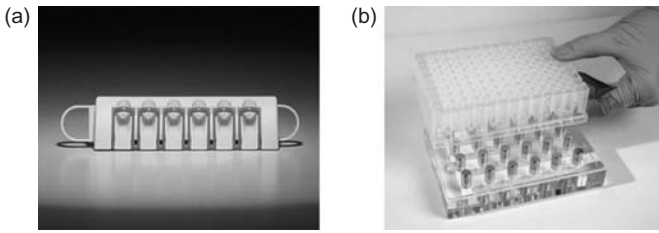


Figure 5.2 Magnetic capture of poly(A)⁺ RNA after base-pairing with oligo(dT) linked to superparamagnetic particles. The beads can be sequestered in (a) microfuge tubes or (b) in 96-well plates. Courtesy of Invitrogen.

there is no need for reprecipitation, which often results in the further loss of unique or low abundance transcripts. The average binding capacity of these beads is 2 μ g poly(A)⁺ RNA per milligram of beads. Transcripts that are affinity purified in this manner are ready for immediate use.

There are numerous protocols that accompany magnetic beads for nucleic acid purification; because protocols are product-specific, they have not been included here. The focus of this chapter is to inform the reader of the available options for improving assay sensitivity. Conveniently enough, there is wide latitude with respect to the selection of starting materials for poly(A)⁺ selection: (a) previously purified total RNA (total cellular or total cytoplasmic); (b) cells growing in culture; or (c) solid tissue. It is very important to realize that in cases where there is a limited amount of starting material it may be prudent to add the beads directly to a whole cell lysate or tissue homogenate for direct harvest of poly(A)⁺ RNA. In the context of assigning gene regulation as either transcriptional or posttranscriptional, the investigator is also cautioned that thymidylate residues recognize and capture *any* transcript with a poly(A) tract, both mRNA *and* hnRNA, and will fail to capture non-poly(A) mRNA. Regardless of the product, the success of the magnetic separation technique requires (1) complete removal of all buffer from the tube containing the magnetic beads before moving on to the next step; and (2) *never* removing anything from the tube containing the magnetic beads unless the tube is in the magnet!

Finally, note that other formats for poly(A)⁺ selection are available from a variety of vendors, including magnetic-based separation products and spin column configurations, all of which in some way target the poly(A) tail. Magnetic separation technology has also been expanded for immunocapture of proteins, cell separation, and other applications.

Oligo(dT)-cellulose column chromatography

Running an “oligo(dT) column” was standard fare in the 1980s and early 1990s. Preparing the matrix and attending to the column was a time-consuming process, the success of which was highly user dependent. The protocol for the in-house

preparation of such a column is presented here; one may also purchase premade oligo(dT)-cellulose spin columns for poly(A)⁺ purification. For example, polyA Spin mRNA Isolation columns (New England BioLabs) are prepacked with microcrystalline oligo(dT) cellulose, each of which can accommodate as much as 1000 μg of total RNA, and centrifugations are performed in a standard microfuge.

For in-house preparation of an affinity column for poly(A)⁺ selection, the oligo(dT) matrix is packed in a column configuration in a 1 ml tuberculin syringe. The eluate is directed through an ultraviolet absorbance detection monitor, such as the UA-6 detector (Teledyne Isco, Inc., Lincoln, NE); this instrument (Fig. 5.3) is also quite popular for a number of protein chromatography applications. As RNA flows from the column and through the optical unit of the instrument (Fig. 5.4) it absorbs UV light in direct proportional to its concentration, as evidenced by the observed tracing on the monitor chart recorder. This method is only practical, however, when beginning with at least several hundred micrograms of total RNA. While this approach to poly(A)⁺ RNA selection is quite a bit more time-consuming than some of the other techniques described here, it has the advantage of allowing the investigator to observe directly the magnitude of the absorbance spectra of the poly(A)⁺ and poly(A)⁻ material. Moreover, the *shape* of the elution profile (Fig. 5.5) is an important indicator of the quality of the sample, the efficiency of elution from the matrix, and the probable utility of the sample in downstream applications. There are two principal drawbacks to this approach, however: (1) the RNA elutes from the column in a relatively large volume, even at slow elution flow rates, making subsequent concentration of the RNA somewhat cumbersome⁶; and (2) due to the lack of sensitivity below 0.1 optical density (OD), this approach is unsuitable for smaller quantities of starting material (less than 200–250 μg total RNA).

Protocol: purification of biophysical quantities of poly(A)⁺ RNA

1. Wear gloves!
2. Prepare oligo(dT)-cellulose according to the instructions of the manufacturer. Typically, 10 mg of oligo(dT)-cellulose is suspended in 3–5 ml of elution buffer (10 mM Tris, pH 7.4; 0.05% SDS⁷) or nuclease-free water.

Note: Depending on the degree of refinement, 1 g of oligo(dT)-cellulose is capable of binding 20–50 OD of poly(A)⁺ RNA. Assuming that 5% of the total RNA sample is polyadenylated, use 3–5 times more oligo(dT)-cellulose material than is required to bind the estimated mass of poly(A)⁺ material in the sample. It is wise to remain within these parameters as excessive matrix bulk may distort the elution characteristics of the

⁶If the concentration of RNA at any time falls below the level of precipitability, it will be necessary to add carrier RNA to accomplish precipitation. Having to resort to this approach defeats the entire purpose of purifying the mRNA in the first place. It is always to one's distinct advantage to keep nucleic acid samples as concentrated as possible.

⁷The inclusion of SDS in elution buffers and binding buffers may not be desirable in some instances because the SDS may precipitate (especially in air-conditioned labs!) and clog the column. Further, residual SDS may inactivate subsequent enzyme activities. Excellent RNase-free technique precludes the requirement for SDS.

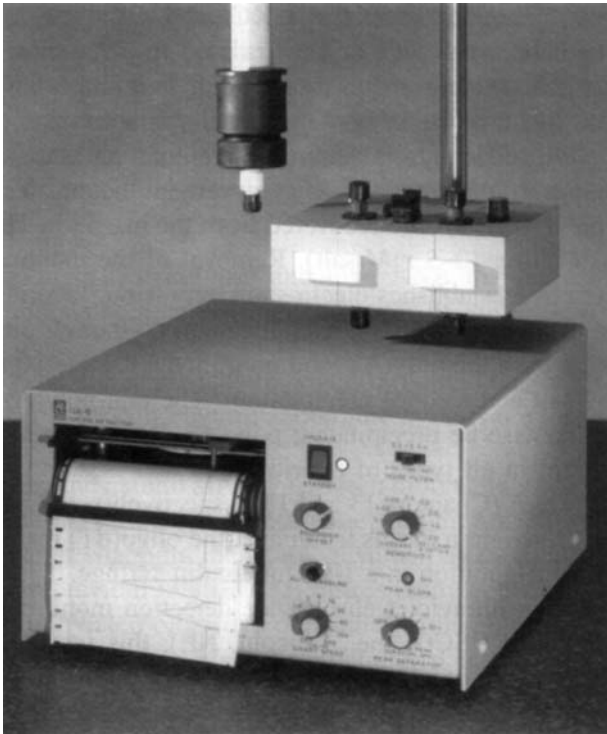


Figure 5.3 UA-6 ultraviolet absorbance detection monitor. Courtesy of ISCO, Inc.

RNA as well as exert pressure on departmental resources: oligo(dT)-celluloses now cost over \$200 per gram! The oligo(dT)-cellulose material is insoluble and should neither be shaken vigorously nor vortexed in an attempt to dissolve it; rather it may be gently resuspended by tapping or gently inverting the tube. For much larger starting quantities of total RNA, as much as 10 mg of RNA can be passed over a larger bore column containing 1 ml of packed oligo(dT)-cellulose.

3. Plug the bottom of a sterile 1 ml tuberculin syringe with autoclaved, shredded polyester fill. Be sure to use only the *smallest amount* necessary to act as a trap for the oligo(dT)-cellulose. Avoid plugging the neck of the syringe; instead, push the polyester down only as far as the bottom of the barrel of the syringe.

Note: Shredded polyester fill is available at most hobby shops and is much easier to manipulate than glass wool. It can be autoclaved briefly (15 min) and should, of course, be handled with gloves. Use an autoclaved long-neck Pasteur pipette to position the polyester at the bottom of the syringe.

4. Apply the oligo(dT)-cellulose to the syringe using an autoclaved 9-inch Pasteur pipette. Allow the matrix to pack itself on top of the polyester. Together, the bed volume of the matrix and the polyester plug should be about 0.1 ml.
5. After the matrix has been set in place, wash with 2–3 ml binding buffer (500 mM NaCl; 10 mM Tris, pH 7.4; 0.05% SDS) to equilibrate the column in high-salt and ensure that the matrix is firmly packed in the column.

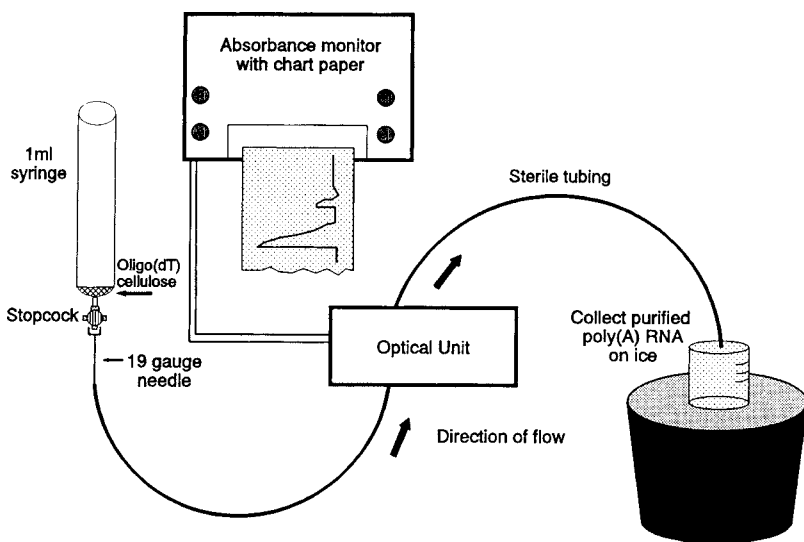


Figure 5.4 Method for purification of large quantities of poly(A)⁺ material. Oligo(dT)-cellulose is prepared and packed in a 1 ml syringe, the flow rate through which is regulated by an autoclavable (RNase-free) stopcock placed beneath. The needle below the stopcock channels the eluate into the tubing which leads to the optical unit of the UV detection monitor. The optical unit is outfitted with a 254 nm absorbance filter for nucleic acid chromatography. The concentration of nucleic acid flowing through the optical unit is interpreted by the magnitude of the deflection of the pen mounted on the monitor, thereby providing the user with a chart paper record of the binding and elution profiles for each sample. The eluate is then directed out of the optical unit and is collected in a sterile tube, usually resting in an ice bucket. The movement of the pen on the chart recorder indicates the precise moment for the investigator to change collecting tubes.

6. Attach a sterile stopcock to the bottom of the syringe in order to regulate the flow rate. Wash the column once again with binding buffer to make sure that no air bubbles have been trapped in or above the stopcock.

Note: Polypropylene stopcocks are quite useful because they can be autoclaved and reused. When assembling the column, it is best to position the stopcock on the bottom of the syringe with the lever in the "open" position so as not to disturb the matrix.

7. Carefully attach a sterile 19-gauge needle or blunted needle to the bottom of the stopcock. Wash the column with binding buffer to make sure that no air bubbles have been trapped in or above the stopcock.
8. Connect the column to an absorbance monitor outfitted with a 254 nm or 280 nm filter, such as the ISCO UA-6 model. (Teledyne Isco, Inc., Lincoln, NE). The connection between the column and the absorbance monitor is made by attaching RNase-free tubing that will accommodate the 19-gauge needle and fit inside the flow cell of the optical unit. The sample enters the flow cell from the bottom and exits from the top: this way, the flow rate of the system is dependent only on hydrostatic pressure.

Note: At this point it is very important to make sure that the column does not run dry. Adjust the flow rate to about 1.5 ml per minute. This relatively slow rate is needed to permit base-pairing between the oligo(dT) matrix and the poly(A) tail to occur.

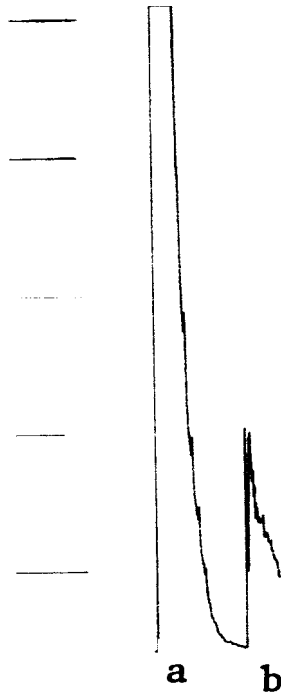


Figure 5.5 Binding and elution profiles of poly(A)⁺ RNA from oligo(dT)-cellulose matrix. The eluate was directed through the optical unit of an ISCO UA-6 ultraviolet absorbance detection monitor. Sensitivity setting was 0.1 OD, which is necessary to observe the elution of less than 100 μg poly(A)⁺ mRNA. (a) Application of 150 μg of total cytoplasmic RNA in binding (high-salt) buffer. The column was washed continuously until the monitor returned to baseline, indicating that all non-poly(A) RNA had passed through the column. (b) Elution of poly(A)⁺ RNA with the application of elution (low-salt) buffer. Degraded samples of RNA tend to produce a wider elution profile without the spiked appearance obvious in this example.

It is not advisable to open and close the stopcock repeatedly because this usually dislodges some of the matrix. Oligo(dT)-cellulose can easily clog the flow cell and result in the loss of valuable material. As with any chromatography, separation is optimal when the column is flowing uniformly.

9. Continue to wash the column with binding buffer until a baseline has been established on the monitor. In this laboratory, the sensitivity is routinely set at 0.1 OD. This sensitivity is necessary to observe the elution of low-to-moderate quantities of polyadenylated RNA (10–200 μg). Of course, the required sensitivity setting is a function of the RNA mass in the sample.

Note: Setting the baseline at the 10% tick is suggested, as differences in the refractive indices of the binding and elution buffers may cause baseline shift.

10. Dissolve the RNA in 500–1000 μl of binding buffer and apply the sample (up to 50 mg RNA per gram of oligo(dT)-cellulose) to the column. Draw the sample into a sterile Pasteur pipette and follow the buffer meniscus down the barrel of the syringe

until the column *almost* runs dry. Carefully release the sample directly onto the matrix and allow it to enter the bed volume. Immediately follow with a copious volume of binding buffer. Do not let the column run dry.

Suggestion: Heat the sample to 65° for 5 min and then rapidly cool it on ice before loading it onto the column. This heating step will disrupt secondary structures that, if present, could make the poly(A) tail inaccessible to the matrix.

11. Observe the enormous deflection of the baseline as the non-poly(A) component of the sample passes through the column (Fig. 5.5a). Continue to wash the column with binding buffer.

Suggestion: Collect and load the eluate onto the column a second time to maximize the capture of the poly(A)⁺ fraction on the matrix. This may be helpful because some poly(A)⁺ RNAs manage to slip through the matrix during the first passage.

12. Continue to wash the column with binding buffer until the monitor pen returns to the baseline. If the poly(A)⁻ fraction is to be saved, it should be collected in a tube on ice as it flows from the optical unit.

Note 1: The poly(A)⁻ fraction may be reprecipitated for future use as a negative control when performing Northern analysis, nuclease protection assay, or a PCR-based assay. Being depleted of poly(A)⁺ RNA, the poly(A)⁻ fraction should fail to produce a signal.

Note 2: Occasionally it becomes necessary to unclog the column if the flow rate diminishes because a sample loaded onto the column contains large amounts of salt and/or genomic DNA carried over in the previous precipitation of the sample. The flow rate may also be diminished by the unexpected appearance of air bubbles. These problems can often be corrected by simply opening the stopcock all the way, to expel air or debris, and then readjusting the flow rate. In more extreme cases, it may be necessary to disrupt the matrix bed directly to unclog it. This approach is acceptable on an as-needed basis. It will usually not cause a loss of resolution because efficient affinity chromatography, unlike gel filtration, is not totally dependent on the geometry of the matrix. Once the oligo(dT)-cellulose particles have settled back in place, one may resume the chromatography. Poly(A)⁺ material will not be eluted from the matrix until the counterion is removed by application of elution buffer.

13. Recover the bound poly(A)⁺ RNA with elution buffer (step 2) prewarmed to 45°. The eluted RNA should be recovered in an RNase-free Corex glass tube, or other appropriate tube, resting in an ice bucket.

Note: Prewarming the elution buffer will greatly facilitate the rapid elution of poly(A)⁺ material in a minimal volume, as it is always desirable to keep nucleic acid samples as concentrated as possible.

14. Reprecipitate the RNA at -20° with 0.1 vol of 3 M sodium acetate, pH 5.2 and 2.5 volumes of ice-cold 95% ethanol. The average recovery per 10⁷ cells is 1–5 µg poly(A)⁺ mRNA.

Note: If the yield is expected to be low or if the eluted RNA is believed to be dilute, reprecipitation of the RNA may be facilitated with the addition of glycogen. Add 10 µl of a 10 mg/ml glycogen solution for every milliliter of poly(A)⁺ eluate collected, followed by the addition of 0.1 vol 3 M sodium acetate, pH 5.2, and 2.5 vol ice-cold 95% ethanol. Store at -20° for at least 4 h.

15. The oligo(dT)-cellulose column can be regenerated by washing it with 2–3 ml of 0.1 N NaOH followed by re-equilibration with binding buffer. A parafilm-sealed column can also be stored under 100–200 µl binding buffer at 4°; don't forget that bacterial cell growth is favored in celluloses stored at room temperature.

Rapid, non-column poly(A)⁺ purification

The selection of poly(A)⁺ RNA can be accomplished without traditional column chromatography. In fact, relatively small masses of starting material (typically less than 200 μg total RNA) mandate a method of separation other than an affinity column. This is mainly because of the dilution of RNA that inevitably occurs and the very low percentage of the total mass that the poly(A)⁺ species represent, not to mention the labor-intensive nature of preparing and running a homemade column. If the yield is extremely small or the sample too dilute, it will be impossible to precipitate the purified RNA without the addition of carrier molecules such as tRNA or glycogen. Of course, such a strategy defeats the purpose of poly(A)⁺ selection in the first place, and excessive glycogen is incompatible with several subsequent RNA applications.

Alternative matrix configurations include the use of magnetic beads, described above, or the use of microcrystalline oligo(dT)-cellulose for the selection of poly(A)⁺ mRNA. Microcrystalline oligo(dT)-cellulose generally can bind in excess of 100 OD (A₂₆₀ units) per gram of matrix. By comparison, the typical binding capacity of the oligo(dT)-cellulose under standard assay conditions generally ranges from 50–80 OD per gram of matrix, depending on the manufacturer, the grade of refinement, and the monovalent cation used to prepare the binding buffer: Li⁺ is better for this application than either K⁺ or Na⁺. The selection of poly(A)⁺ RNA is best accomplished by repeated suspension of the matrix in traditional high-salt buffers, followed by centrifugation. Using this approach, one avoids the excessively slow flow rates commonly associated with column chromatography, and it should be noted that microcrystalline oligo(dT)-cellulose is not recommended for use in column format unless pre-packed by the manufacturer. If scaled down, the entire procedure can be carried out in a microfuge tube. The eluted RNA thus remains in a relatively small volume and at a concentration that permits precipitation with salt and alcohol alone, without an added carrier.

The following protocol is our variation of one of several procedures for microcrystalline oligo(dT)-cellulose. Although this type of separation can also be used with standard oligo(dT)-celluloses, the microcrystalline grade maximizes quantitative recovery. Note that all requirements for controlling RNase activity must be satisfied for quantitative and qualitative recovery.

Protocol: non-column poly(A)⁺ purification

1. Wear gloves!
2. Weigh out an appropriate mass of matrix material. For microcrystalline oligo(dT)-cellulose, use 100 mg (dry weight) for each 150 μg of total RNA starting material. Place matrix in a nuclease-free microcentrifuge tube.

Note: Oligo (dT)-cellulose should be weighed out under RNase-free conditions. While autoclaved or baked metal spatulas may be used (see Chapter 7 for details), it is better to gently tap the bottle of oligo(dT)-cellulose to remove a suitable mass.

3. Add 500 μ l binding buffer (500 mM LiCl; 20 mM Tris-Cl, pH 7.5; 1 mM EDTA; optional 0.05% SDS) to the oligo(dT)-cellulose. Allow the beads to equilibrate in this high-salt buffer for 5 min.
4. Centrifuge the tube containing the equilibrated oligo(dT)-cellulose at 200 \times g for no more than 30 s. Carefully remove and discard the supernatant with disturbing the oligo(dT)-cellulose pellet. Set tube aside.
5. In a separate tube, dissolve RNA in nuclease-free H₂O in a final volume of 100 μ l and heat to 65° for 5 min.

Note: The heating step in this and related protocols is necessary to denature RNA molecules, thereby making the poly(A) structure more available for base-pairing to the matrix.

6. Add an equal volume of binding buffer to the heat-denatured RNA and allow it to cool to room temperature. Add the mixture to the oligo(dT)-cellulose that was set aside in step 4.
7. Place the tube in a microfuge tube rack and allow the sample to stand at room temperature for 5 min. Agitate the tube gently by hand once or twice during this brief incubation.
8. Centrifuge at 1500 \times g for 5 min, at 4° if possible. *Carefully* transfer the supernatant to a fresh tube and label as “poly(A)⁻ RNA”. Set tube aside on ice.

Note 1: If SDS is added to the binding buffer, carry out the centrifugation at room temperature or only as low as 15°. Otherwise, SDS will begin to precipitate in the tube.

Note 2: The poly(A)⁻ fraction can be extremely useful as a negative control in a variety of applications or can itself be used as carrier RNA in future experiments.

9. Wash the oligo(dT)-cellulose up to five times with 500 μ l aliquots of binding buffer. Centrifuge at 1500 \times g after each wash and then *carefully* remove the supernatant between each wash.
10. Elute the poly(A)⁺ RNA from the matrix with one or more 50 μ l aliquots of elution buffer (10 mM Tris-Cl, pH 7.5; 1 mM EDTA; optional 0.05% SDS). Alternatively, RNA may be eluted in nuclease-free H₂O.

Note: The number of elution buffer aliquots needed to recover all of the bound material will be a function of poly(A)⁺ mass and the volume of elution buffer used. The first aliquot of elution buffer will contain a great majority of the poly(A)⁺ material. Logistically speaking, it may be useful to elute the poly(A)⁺ in empirically-determined smaller volumes. As needed, the RNA should be reprecipitated in one or more microfuge tubes. This is easily accomplished by adding 0.1 vol 3 M NaOAc (pH 5.2) and 2.5 volumes ice cold 95% ethanol. Store sample(s) on dry ice for 20 min or at -20° for at least 4 h.

11. Following precipitation and centrifugation, carefully decant the ethanol supernatant. Wash the RNA pellet two or three times with 500 μ l aliquots of 70% ethanol (prepared in sterile H₂O). A final wash with 95% ethanol may facilitate air drying. Resuspend RNA in a minimum volume of TE buffer, pH 7.5 or sterile H₂O.

12. Determine poly(A)⁺ concentration by measuring A₂₆₀, as described in Chapter 6.

13. Aliquot rehydrated RNA and store at -80° until further use.

References

- Aviv, H. and P. Leder. (1972). Purification of biologically active globin messenger RNA on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci.* **69**, 1408-1412.
- Aziz, N. and H.N. Munro. (1986). Both subunits of rat liver ferritin are regulated at a translational level by iron induction. *Nucleic Acids Res.* **14**, 915-927.

- Buell, G.N., M.P. Wickens, F. Payvar and R.T. Schimke. (1978). Synthesis of full length cDNAs from four partially purified oviduct mRNAs. *J. Biol. Chem.* **253**, 2471–2482.
- Cole, C.N. and T.P. Stacy. (1985). Identification of sequences in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation. *Mol. Cell. Biol.* **5**, 2104–2113.
- Conway, L. and M. Wickens. (1985). A sequence downstream of A-A-U-A-A is required for formation of simian virus 40 late mRNA 3' termini in frog oocytes. *Proc. Natl. Acad. Sci.* **82**, 3949–3953.
- Darnell, Jr., J.E. (1979). Transcription units for mRNA production in eukaryotic cells and their DNA viruses. *Prog. Nucleic Acids Res.* **22**, 327–353.
- Efstratiadis, A., F.C. Kafayos, A.M. Maxam and T. Maniatis. (1976). Enzymatic *in vitro* synthesis of globin genes. *Cell* **7**, 279–288.
- Gil, A. and N.J. Proudfoot. (1984). A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. *Nature* **312**, 473–474.
- Gilham, P.T. (1964). Synthesis of polynucleotide celluloses and their use in fractionation of polynucleotides. *J. Am. Chem. Soc.* **86**, 4982–4985.
- Kramer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *Ann. Rev. Biochem.* **65**, 367–410.
- Manley, J.L. (1988). Polyadenylation of mRNA precursors. *Biochem. Biophys. Acta* **950**, 1–12.
- McDevitt, M.A., M.J. Imperiale, H. Ali and J.R. Nevins. (1984). Requirement for a downstream sequence for generation of a poly(A) addition site. *Cell* **37**, 993–999.
- Mercer, J.F.B. and H. Naora. (1975). A comparison of the chromatographic properties of various polyadenylate binding materials. *J. Chromatogr.* **114**, 115–128.
- Meyuhas, O., Y. Bibrerman, P. Pierandrei-Amaldi and F. Amaldi. (1996). Isolation and Analysis of Polysomal RNA. In *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis* (P.A. Krieg Ed.). Wiley Liss, Inc, New York. pp. 65–81.
- Meyuhas, O., A.E. Thompson and R.P. Perry. (1987). Glucocorticoids selectively inhibit the translation of ribosomal protein mRNAs in P1798 lymphosarcoma cells. *Mol. Cell. Biol.* **7**, 2691–2699.
- Moore, C.L. and P.A. Sharp. (1985). Accurate cleavage and polyadenylation of exogenous RNA substrate. *Cell* **41**, 845–855.
- Nevins, J.R. and J.E. Darnell. (1978). Steps in the processing of Ad2 mRNA: poly(A)⁺ nuclear sequences are conserved and polyadenylation precedes splicing. *Cell* **15**, 1477–1493.
- Pierandrei-Amaldi, P. and F. Amaldi. (1994). Aspects of regulation of ribosomal protein synthesis in *Xenopus laevis*. *Genetica* **94**, 181–193.
- Proudfoot, N.J. and G.G. Brownlee. (1976). 3' non-coding region sequences in eukaryotic mRNA. *Nature* **263**, 211–214.
- Ross, J. (1995). mRNA stability in mammalian cells. *Microbiol. Rev.* **59**, 423–450.
- Sadofsky, M. and J.C. Alwine. (1984). Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. *Mol. Cell Biol.* **4**, 1460–1468.
- Takagaki, Y., L. Ryner and J.L. Manley. (1988). Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* **52**, 731–742.
- Williams, J.G. (1981). The preparation and screening of a cDNA clone bank. In "Genetic Engineering." (R. Williamson, Ed.), Vol. 1. Academic Press, New York. pp. 1–59.
- Ziff, E.B. (1980). Transcription and RNA processing in the DNA tumor viruses. *Nature* **287**, 491–499.

This page intentionally left blank

6 Quality Control for RNA Preparations

Rationale	139	Protocol	149
Quality control technique 1: UV spectrophotometry and absorption ratios	140	Quality control technique 3: UV shadowing	150
Spectrophotometric methods	140	Protocol	151
Determination of nucleic acid concentration	141	Quality control technique 4: Sample capacity to support RT-PCR	152
Determination of nucleic acid purity	143	Quality control technique 5: Northern analysis	153
Non-spectrophotometric methods	145	Quality control technique 6: Sample capacity to support <i>in vitro</i> translation	153
Quality control technique 2: Electrophoretic profile of the RNA	147	References	154

Rationale

Every laboratory application involving purified RNA requires a quality assessment of the sample, particularly the mRNA fraction. Because of the labor-intensive and costly nature of most downstream assays, it is prudent to show that the template material is intact so as to maximize productivity, reproducibility, and sensitivity. One would certainly not begin with degraded or undetectable quantities of RNA and simply “hope for the best” for the same reasons that one would not attempt to build a house with wood that had rotted.

Checking a purified sample of RNA for integrity and overall quality may be accomplished in a variety of ways, some of which are also very good positive-control techniques for a number of transcription-based assays. Of the methods delineated below, spectrophotometric analysis and examination of the electrophoretic profile of each sample (quality control techniques 1 and 2, respectively) should always be performed shortly after purification and, if the RNA has been stored for a while, a second aliquot should be tested just prior to use. Do not ever assume that a sample will be handsome-looking because a kit was used for purification or that a sample previously characterized remains intact after storage for several months, even at -80° .

Quality control technique 1: UV spectrophotometry and absorption ratios

Nucleic acid concentration can be determined with minimal fanfare by spectrophotometric analysis. The absorbance of ultraviolet (UV) light provides a direct means of determining both nucleic acid concentration and purity, as well as an indirect means of calculating the total yield from the biological source and even an estimation of the mass of RNA per cell, assuming that the approximate cell number is known at the onset of the isolation.

Spectrophotometric methods

Measuring the UV absorption analysis is the easiest, most rapid method for determining yield, purity, and probable utility in downstream applications, and the importance of making these measurements cannot be overstated. Pure samples of nucleic acid have a characteristic absorbance profile between 230 and 320nm (Fig. 6.1). Deviations from the standard shape of the curve indicate the presence of contaminants, the nature of which is suggested by the location of the skewing of the curve. In addition, reaction efficiency is heavily dependent on the input mass and sample purity. For example, precise enzyme:primer:RNA template ratios are critically important for efficient conversion of RNA into cDNA, so one needs to know the concentration of the template material with great accuracy. Further, looking for up- or down-regulation of specific transcripts requires some basis for normalization of samples to each other. This normalization is performed, first and foremost, at the level of RNA input mass.

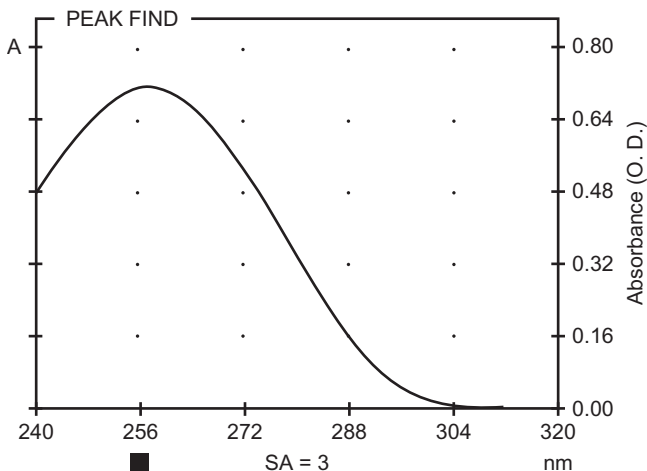


Figure 6.1 Typical UV absorbance spectrum of purified RNA. The spectrophotometric profile for purified DNA is similar. Note the positive slope of the curve below 260 nm, and the negative slope above 260 nm.

Determination of nucleic acid concentration

The concentration and purity of RNA and DNA samples are determined readily by taking advantage of the ability of nucleic acids to absorb UV light, with an absorbance maximum at 260 nm. Absorbance measurements at this wavelength permit the direct calculation of nucleic acid concentration in a sample:

$$[\text{RNA}]\mu\text{g/ml} = A_{260} \times \text{dilution} \times 40$$

where

A_{260} = absorbance, in optical densities, at 260 nm (OD_{260})

Dilution = dilution factor

40 = average extinction coefficient of RNA ($40\mu\text{g}/\text{OD}_{260}$)

The concentration of a sample of double-stranded DNA (dsDNA) can be determined in a similar manner:

$$[\text{dsDNA}]\mu\text{g/ml} = A_{260} \times \text{dilution} \times 50$$

where

A_{260} = absorbance, in optical densities, at 260 nm (OD_{260})

Dilution = dilution factor

50 = average extinction coefficient of dsDNA ($50\mu\text{g}/\text{OD}_{260}$)

The concentration of single-stranded DNA (ssDNA) such as oligonucleotides can also be determined easily:

$$[\text{ssDNA}]\mu\text{g/ml} = A_{260} \times \text{dilution} \times 33^*$$

where

A_{260} = absorbance, in optical densities, at 260 nm (OD_{260})

Dilution = dilution factor

33 = average extinction coefficient of ssDNA ($33\mu\text{g}/\text{OD}_{260}$)

Note that the above equations render concentration data in units of $\mu\text{g/ml}$; for the molecular biologist, the mathematics associated with sample dilution and reaction assembly are often greatly simplified when concentration is expressed in $\mu\text{g}/\mu\text{l}$, so remember to divide by 1000. Also, the units of OD_{260} are ml cm^{-1} , which is useful to know if converting between mass and molarity (see Appendix B for clarification, as needed).

For greater precision in determining nucleic acid concentration, the extinction coefficient for a particular (unique) RNA or DNA sample can be calculated,

*The extinction coefficient for each oligonucleotide is unique and is always reported on the product "spec sheet" from the manufacturer that accompanies the product. If the precise extinction coefficient for a particular oligonucleotide is unknown, use a value of $33\mu\text{g}/\text{OD}_{260}$.

using Beer's Law¹, if the G + C content of the organism is known. The average extinction coefficients of 40, 50, and 33 for RNA, dsDNA, and ssDNA, respectively, are more or less the standards for most molecular biology applications and are preprogrammed into later model spectrophotometers for easy use. Since the extinction coefficients of thymine and uracil differ, however, DNA contaminating RNA samples and RNA contaminating DNA samples can falsely elevate or depress spectrophotometric readings. DNA can be removed from an RNA sample by incubation in RNase-free DNase I. This technique is described in detail in Appendix F. Similarly, incubation of a DNA sample with RNase A, followed by reprecipitation, should remove virtually all contaminating RNA from the sample, as described in Appendix G. In addition, when contaminating RNA is present, DNA molecules often display increased electrophoretic mobility, a phenomenon that may well result in an underestimation of DNA fragment size or size range.

For older-style spectrophotometers, UV measurements are best taken by diluting the sample 1:200–1:500 in acid-washed quartz cuvettes². Alternatively, a new generation of disposable plastic cuvettes is now manufactured that permits measurements in the UV portion of the spectrum. The reader is cautioned that not all plastic cuvettes are UV-rated, and one should double-check the manufacturer's description of the product before ordering. Ideally, UV-rated plastic cuvettes should be used only once, but certainly not more than 3–4 times, after which the plastic is changed by exposure to UV light and the data become wholly unreliable.

When determining nucleic acid concentration, one should be certain to use a sufficient mass to obtain at least the minimum absorbance recommended by the manufacturer of the spectrophotometer. For example, spectrophotometric measurements of less than 0.1 optical density (OD) are often unreliable in older instrumentation. Larger-volume cuvettes require a minimum volume to be readable and, when used, the increased volume may have the unfortunate consequence of diluting the sample below the level of detection. In cases such as this, the investigator is urged to invest in smaller-volume, though more expensive, quartz cuvettes. It is also important to note that these measurements reflect the total mass of nucleic acid in a sample, and not the relative contribution of any single component, such as mRNA in a mixture of total cellular RNA.

One of the more amazing improvements in spectrophotometric instrumentation is the NanoDrop™ product line (NanoDrop Technologies, Wilmington, DE). These fiber-optic-based instruments are ideal for processing several samples in

¹According to the Bouguer–Beer Law, commonly known as Beer's Law, $A = \epsilon bc$ where A is absorbance, ϵ is the molar extinction coefficient, b is the light path length (1 cm is standard), and c is the concentration of the absorbing material, such as adenosine, cytidine, guanosine, thymidine, and uridine. This equation can be rearranged such that $\epsilon = A/bc$. In a mixture of polynucleotides, the extinction coefficient can be determined because $A_{TOT} = A_1 + A_2 + A_3 \dots$. For example, a precise extinction coefficient for human RNA is 44.19.

²Quartz cuvettes can be efficiently cleaned and purged of ribonuclease activity by soaking them in chromic acid and then rinsing them thoroughly with sterile (nuclease-free) water. Exercise extreme caution when handling chromic acid.

rapid succession and are capable of full-spectrum visible and UV wavelength measurements. As shown in Fig. 6.2, as little as 1 μ l of sample is required for rapid, precision profiling of the sample. In this laboratory we routinely load 2 μ l per sample directly onto the pedestal, and there are no required consumables, except for the micropipette tip. Perhaps best of all, there are no cuvettes; the sample is retained by surface tension alone while the absorption measurements are performed. The NanoDrop instrument can also be used to measure protein concentration and fluorescent dyes. An alternative fluorometer platform is also available for related spectrophotometric applications.

Determination of nucleic acid purity

Calculations based on absorbance at 260 nm provide little information about the quality and purity of the sample; in the presence of excess salt, contaminating proteins and/or carryover organic solvents, the absorbance value can be skewed significantly. For this reason, calculation of the so-called “260 to 280 ratio” (A_{260}/A_{280}), the “260 to 230 ratio” (A_{260}/A_{230}), and the “260 to 240 ratio” (A_{260}/A_{240}) provide a reasonable estimate of the purity of the preparation (Table 6.1). A pure sample of RNA has an A_{260}/A_{280} ratio of 2.0 ± 0.1 , and a pure sample of DNA has an A_{260}/A_{280} ratio of 1.8 ± 0.1 . For both RNA and DNA, the A_{260}/A_{230} should be greater than 2.0 and less than 2.4. Variations outside this range generally indicate contaminants, and investigators attuned to this detail may wish to take corrective action. A low A_{260}/A_{280} usually indicates a protein contamination that carried over during the RNA isolation. A low A_{260}/A_{230} may be caused by carry-over guanidinium (from chaotropic lysis buffer) or β -mercaptoethanol, and this is much more of a problem than most people think. One method for removing excess guanidinium buffer is to reprecipitate the RNA with sodium acetate and ethanol, followed by washing the pellet 2–3 times with 70% ethanol. This same procedure should be followed if the A_{260}/A_{230} is greater than 2.4. When working with plant RNA, a low A_{260}/A_{230} may indicate polysaccharide contamination. Finally, use the A_{260}/A_{240} ratio, which should be around 1.4 for pure nucleic acids, to detect

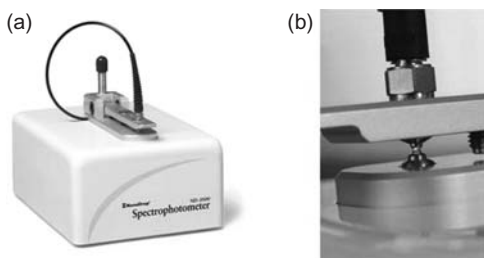


Figure 6.2 NanoDrop spectrophotometer. (a). The very small foot print of this instrument is very convenient for placing it almost anywhere in the lab. (b). Surface tension holds a 1–2 μ l aliquot of a sample in place between the pedestal and the arm while the fiber optics system measures concentration, displays the absorbance profile, and calculates standard ratios for assessing purity. Courtesy of NanoDrop, Inc.

Table 6.1 Strategic Indicators of Nucleic Acid Purity

	DNA	RNA	Problem ratios
A_{260}/A_{280}	1.8	2.0	Ratios below 1.8 suggest protein contamination
A_{260}/A_{230}	2.0–2.4	2.0–2.4	Ratios below 1.8 suggest organic compound contamination
A_{260}/A_{240}	1.4	1.4	Ratios below 1.4 suggest an excessive amount of salt in the sample

excess amounts of salt. When A_{260}/A_{240} is less than 1.4, one should consider reprecipitation of the sample followed by multiple washes with 70% ethanol, and a final 95% wash to accelerate drying the pellet.

On several occasions over the last few years the accuracy of the using A_{260} absorbance as an indicator of concentration and A_{260}/A_{280} as a purity indicator have been called into question (Glasel, 1995; Huberman, 1995; Manchester, 1995; Manchester, 1996). The pH of the diluent does, in fact influence the absorbance ratio, and is most reliable when measurements are made between pH 7.5–8.5 (Wilfinger *et al.*, 1997). It is now commonplace to dilute both DNA and RNA samples in TE- or phosphate buffer to establish a pH within this range. The take-home lesson is that biochemical quality water, which is usually slightly acidic with pH around 6.0, is not the best choice for accurate measurement of the concentration of nucleic acid in a sample.

Another common problem associated with the isolation of RNA is the incomplete removal of protein from the cell lysate. Because proteins increase the absorbance at 280 nm, the resulting decrease in the A_{260}/A_{280} ratio is a reliable indicator of a purity problem with a sample. In the event of incomplete protein extraction, the sample (RNA or DNA) should be dissolved 100–200 μ l sterile (nuclease-free) buffer, extracted once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and then once with an equal volume of chloroform or chloroform:isoamyl alcohol (24:1). The appearance of a protein interphase, no matter how small, between the lower organic phase and the upper aqueous phase clearly indicates protein contamination. RNA samples in large volumes may then be reprecipitated to increase the concentration of the sample. This type of manipulation should not be used gratuitously, however, as the stability and recovery of the sample is compromised by repeated precipitation, not to mention the tremendous opportunity to introduce nuclease activity into the sample. Performing one extra phenol:chloroform extraction to remove contaminating protein *during* the isolation procedure is definitely worthwhile if an additional extraction appears to be warranted and may well preclude a lot of aggravation later. Precipitated samples of RNA should be washed with 70–75% ethanol and the concentration determined as described earlier.

Although the value obtained for a particular sample is a good indicator of the quality of the sample, it is also of paramount importance to examine the absorbance profile of the sample, usually between 230 and 320 nm. Pure samples

of nucleic acids will produce a characteristic skewed bell curve (Fig. 6.1), which exhibits an absorbance maximum close to 260 nm. The absorbance of ultraviolet light by nucleic acids is 0 at 320 nm, and below 230 nm the absorbance profile has little bearing on computation of the concentration and determination of the purity of a sample. High quality nucleic acid samples manifest an absorption profile with a positive slope between 230–260 nm and a negative slope above 260 nm. This is not observable when measuring absorbance only at 260 nm and 280 nm.

The rationale for doing these types of preliminary analyses to gain information about the quality of a sample should be obvious. In the absence of UV spectrophotometry, the amount of nucleic acid in a sample would be largely unknown, thereby making normalization among samples a virtual impossibility. Without computation of the various absorbance ratios, it may not be immediately obvious if contaminants have co-purified with the sample. In the absence of a wavelength scan between 230 and 320 nm, there will be no indication of the shape of the curve. It is frequently noted in the laboratory that contaminants change the characteristic shape of the curve at wavelengths other than 260 nm or 280 nm. For example, excessive quantities of carryover salt (nucleic acid co-precipitate) can mask the positive slope of the curve that is usually associated with the 240–260 nm portion of the spectrum. Washing a nucleic acid co-precipitate with 70–75% ethanol immediately following centrifugation (as suggested in several protocols) will remove most of the salt³ that was used to drive the precipitation of the nucleic acid molecules in the first place. In other cases, investigators have been known to remove excess salt by sample dialysis. In the absence of a wavelength scan, the presence of excess salt, and probable negative impact on downstream applications, would not be evident.

Non-spectrophotometric methods

In the absence of fairly expensive spectrophotometric equipment, or if the amount of nucleic acid in a sample is very small, there are other rather unsophisticated methods through which a reasonable attempt can be made to assess the concentration of nucleic acid in a sample. First, the prevalence of image analysis software in the molecular biology laboratory affords an opportunity for quantifying nucleic acid concentration by digitizing a fluorescence image of a sample that has been run out on a gel and comparing it to a known mass standard. One may purchase mass standards from any of a number of suppliers. Dilutions of the commonly used molecular weight standard Φ X174 or λ DNA can be electrophoresed side by side on a gel along with RNA or DNA of unknown concentration (Fig. 6.3). Image analysis software can then count the pixels that make up the image of all bands and smears in a lane, and then compare the value for each unknown sample against a mass standard. In this lab, this approach for determining nucleic acid concentration is often used to assess

³Sodium acetate (NaOAc) is particularly easy to remove with 70% ethanol washes.

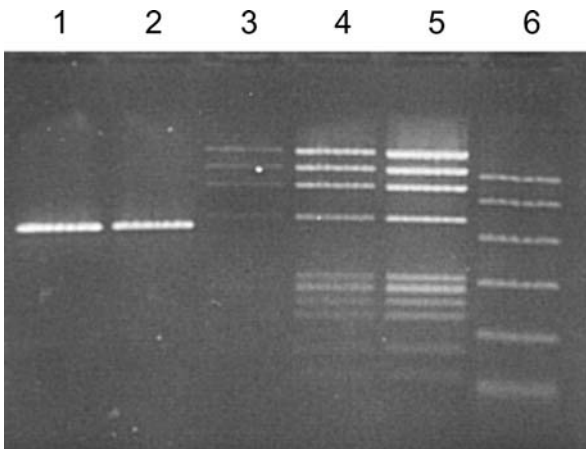


Figure 6.3 Determination of sample mass and molecular weight by comparison with dilutions of a DNA mass standard. Electrophoresis was performed on a 2.5% agarose gel. (1) 2 μ l sample of unknown concentration. (2) 1 μ l sample of unknown concentration. (3, 4, 5) 100 ng, 200 ng, and 400 ng, respectively, of Φ X174 DNA. (6) PCR Marker (Promega); molecular weights are 1000 bp, 750 bp, 500 bp, 300 bp, 150 bp, and 50 bp. Calculation method and further details are found in Chapter 11.

the efficiency of cDNA synthesis by quantifying the smear. The method is rapid and usually quite accurate, especially in view of the ready availability of super-sensitive stains such as SYBR Green, Gel Star, and SYBR Gold. The subject of image analysis is discussed and critiqued in detail in Chapter 10.

Second, several companies sell kits for fluorescence-based determination of nucleic acid concentration. One mixes a fluorescent compound with the sample, which then fluoresces in proportion to the mass of nucleic acid with which it has bound. This procedure requires a fluorescence detector for making fluorescence measurements and also requires that the user prepare a standard curve against known samples in order to make precision concentration determinations.

A third, non-spectrophotometric determination of nucleic acid concentration involves making an estimate by simple visual inspection. The method involves spotting dilutions of known amounts of a nucleic acid and samples of unknown concentration on to an agarose plate containing a fluorescent dye. Upon UV irradiation, the dye will fluoresce in proportion to the amount of nucleic acid present, and one may make a *rough estimate* of the amount of nucleic acid in the unknown by comparing fluorescence intensity between the known and unknown samples. Performing this somewhat imprecise assay require the preparation of agarose plates containing either 0.5 μ g/ml ethidium bromide or 1X SYBR Green⁴.

⁴To prepare ethidium bromide/agarose plates, prepare a solution of 1% agarose in water, cool to 55°, and then add ethidium bromide stock solution (10 mg/ml) to a final concentration of 0.5 μ g/ml. Carefully swirl the flask to mix, and avoid inhaling vapors. The molten agarose/ethidium bromide mixture can then be poured into 100mm Petri dishes and allowed to solidify. SYBR Green plates are prepared in the same manner, substituting 1X SYBR Green in place of the ethidium bromide.

Dilutions of the standards and the test material are best applied in 5 μ l aliquots so as to avoid sample diffusion, which would make quantitation even more difficult. These plates can be observed directly on the surface of a UV transilluminator or irradiated from above with a handheld UV monitor.

CAUTION: *Handheld UV irradiating devices can inflict injuries as serious as those from transilluminators. Be sure to wear proper eye protection and cover exposed skin when working with UV light.* Be sure to allow samples to be absorbed into the agarose before assessing the fluorescence intensity of the ethidium bromide. Alternatively, dilutions of a nucleic acid sample of unknown concentration can be added to an equal volume (5 μ l or less) of a 1 μ g/ml solution of ethidium bromide or 1X solution of SYBR Green. These samples can then be applied directly onto plastic wrap, and irradiated with UV light, and the resulting fluorescence can be compared to DNA standards of known concentration (Wienand *et al.*, 1978). Although this method is barbaric, it may be a useful teaching tool when funds are limited.

Quality control technique 2: Electrophoretic profile of the RNA

In general, running out an RNA sample on an agarose gel is the single best diagnostic that the investigator has at his disposal; parameters and protocols for electrophoresis are described both in Chapter 9 and in Appendix M. Chemically intact, biologically competent RNA always produces a characteristic banding profile when denatured, which can be observed simply by heating the sample briefly before loading it onto an agarose gel. Along these same lines, one may incubate a high-quality RNA sample with a solution to test for the presence of nuclease activity. If, after incubation, the RNA shows any signs of degradation, then the solution being tested is probably tainted with RNase.

There is no better indicator of the probable utility of an RNA sample than measurement of the fluorescence ratio between the 28S rRNA and the 18S rRNA. For mammalian RNA, a minimum fluorescence ratio of 2:1 is desirable and a ratio of 2.5:1 or greater is better yet. It has been the experience of this Author and many others that the smaller the 28S:18S ratio, the less likely the sample will yield representative data or do so in a reproducible manner. One may simply “eyeball” the gel to estimate the 28S:18S ratio or, as needed, the measurement can be performed using image analysis software. For higher eukaryotes, the larger the ratio, the better. Be aware, however, that other aspects of the electrophoretic profile are also of key importance.

Excellent RNA (eRNA) samples show a minimum amount of smearing above, between, and below the 28S and 18S rRNAs. Transfer RNA (tRNA), and the low-molecular-weight 5S and 5.8S rRNA species all co-migrate at the leading edge of the gel and, when stained, usually appear as an indistinct splotch in the vicinity of the bromophenol blue tracking dye that is often part

of standard gel loading buffers. Lack of definition to the 28S and 18S rRNAs usually means that the sample has been subjected to nuclease assault, especially if the smearing is confined to the lower portion of the gel, below the level of the 18S rRNA, which can be observed in intact RNA in adjacent lanes on the same gel. Heavy smearing along the length of the gel may indicate limited degradation of the sample, though it is much more probable that the RNA was not denatured properly prior to electrophoresis: the resultant smear may very well reflect persistent secondary RNA structure. Further, this Author has also observed situations in which the RNA was not completely dissolved, resulting in a characteristic smear along the length of the lane, as described in Chapter 9. While it is desirable to keep nucleic acid stocks concentrated, as described above, it is also necessary to make sure that the entire RNA or DNA sample is completely dissolved so that each aliquot removed from the stock tube is a representative aliquot. Detergents and excess salt in the sample can also cause smearing of the RNA and, in some cases, inhibit the entry of the RNA into the gel itself. If this is a chronic difficulty, the use of *N*-laurylsarcosine (sarkosyl) may alleviate detergent-related symptoms. Moreover, if care was not taken to remove DNA⁵ from an RNA prep, as described Chapter 2 and Appendix F, contaminating DNA will coprecipitate with the RNA in the final stages of RNA purification, reeking havoc in downstream applications.

The inclusion of formaldehyde and formamide, necessary when performing the complete Northern analysis, are optional at this stage, and the ribosomal RNAs are usually quite evident when electrophoresed under non-denaturing conditions. It should be noted that the inclusion of 2 M urea will facilitate sample denaturation, producing a gel that is easier to interpret, compared to the omission of these denaturants. Following a relatively brief period of non-denaturing electrophoresis, visual inspection of the predominant 28S and 18S ribosomal RNA (rRNA) species confirms that the RNA is intact (Fig. 6.4).

Genomic DNA on an EtBr- or SYBR Green-stained RNA gel appears as an area of fluorescence within and just below the well into which the sample was loaded (Fig. 6.5). If this phenomenon is observed, it is wise to ascertain the cause of the DNA carryover before using the RNA for any type of quantitative analysis. Of course, it is also critical to understand that failure to observe high-molecular-weight fluorescence does not mean that the RNA sample is devoid of DNA, only that if present the DNA is below the level of detection. Needless to say, contaminating DNA has a profound negative influence on the outcome of most RNA-based PCR assays in its capacity as an alternative template for primers. The only reliable methods to ensure the absence of DNA from RNA samples are to incubate the sample with RNase-free DNase *and* run a control reaction that will produce a result only if genomic DNA is present in the sample, for example, a “no reverse transcriptase” control, as described in Chapter 18.

⁵Techniques for the removal of DNA from RNA preparations include differential centrifugation of intact nuclei, treatment of the sample with RNase-free DNase I, repeated extraction with acid-phenol, isopycnic partitioning of the RNA, or any combination thereof.

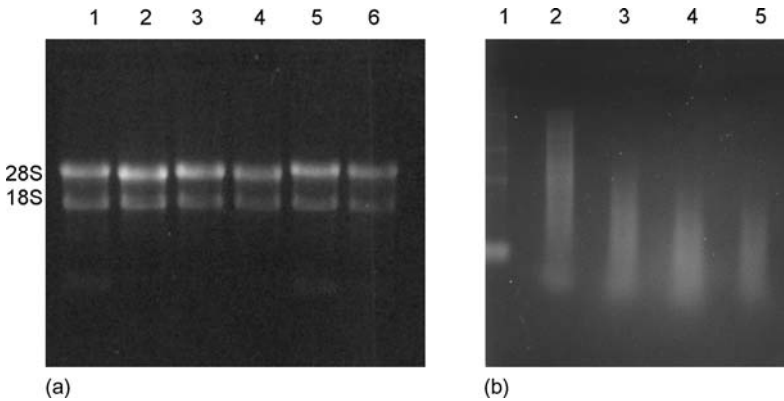


Figure 6.4 Assessment of the integrity of RNA samples. (a) Minigel (7 cm \times 8 cm) electrophoresis of RNA isolated Swiss mouse 3T3 fibroblasts. Five micrograms of purified RNA was electrophoresed in a 1.2% agarose-formaldehyde gel for 1.5 h at 50V and then stained with SYBR Green II. The clear definition of the ribosomal 28S and 18S rRNA species and the greater fluorescence of the 28S rRNA in lanes 1–6 demonstrates the integrity of the samples. (b) Going, going, gone. Minigel (6 cm \times 9 cm) electrophoresis of human fibroblast RNA. A representative aliquot from four different samples was electrophoresed in a 1.2% agarose-formaldehyde gel for 1.5 h at 50V and then stained with ethidium bromide. The RNA molecular weight standard is barely visible in lane 1. Samples in lanes 2–5 show increasing degrees of degradation, most likely due to RNase contamination during the isolation procedure. This is an excellent example of what not to do.

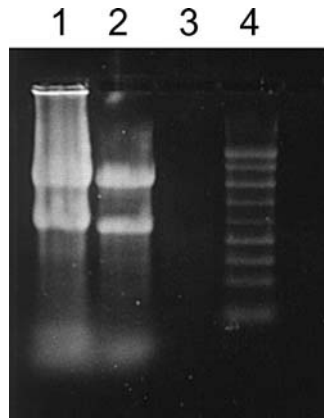


Figure 6.5 SYBR Green II staining of RNA samples. Fluorescence confined to a well (lane 1) into which a sample was loaded strongly suggests genomic DNA contamination.

Protocol

The integrity of experimental RNA can be determined by minigel (gel dimensions: 7 cm \times 8 cm) electrophoresis through a standard, non-denaturing 1.2% agarose gel. The RNA sample itself is briefly heat-denatured just prior to electrophoresis in order to disrupt secondary structure.

1. Prepare a 1.2% solution of molten agarose in $1 \times$ TAE⁶ buffer. Cast gel when agarose has cooled to 65°.

Note: If desired, EtBr can be added directly to the molten agarose to a final concentration of 0.5 µg/ml. In so doing, the gel can be examined as soon as the electrophoresis has been completed. Do not add SYBR Green to gels prior to electrophoresis because it will profoundly interfere with the formation of sharp bands. Always stain gels with SYBR Green after electrophoresis.

2. While the gel is solidifying, dilute 1 to 5 µg RNA in 10 µl sterile H₂O or TE buffer.

Note: If the RNA proves difficult to denature by heating alone, then add 2.5 µl formaldehyde (37% stock solution) and 5.0 µl formamide to the 10 µl aliquot of RNA. Handle formaldehyde and formamide cautiously and in a chemical fume hood, according to the manufacturers Material Safety Data Sheet (MSDS). Alternatively, 2 M urea can be added to the sample.

3. Heat to 65° for 5 min. Pulse centrifuge to collect sample at bottom of tube.

4. Add 1 µl of 10× loading buffer (50% glycerol, 1 mM EDTA, 0.25% bromophenol blue) and then load the gel.

Note 1: Use 2 µl loading buffer if formaldehyde and formamide were added to the RNA sample.

Note 2: A molecular weight marker is not needed; this is a quick and dirty assessment of the integrity of the sample.

5. Electrophorese at 75 to 100 volts until dye front has migrated 2 to 3 cm into the gel. This is all that is necessary to visualize the sample; however, allowing the gel to run longer will certainly favor greater resolution.

6. Unless the dye was added directly to the agarose, stain gel for 10 min in a solution of EtBr (0.5 µg/ml in H₂O). Alternatively, stain the gel in SYBR Green II (1× in 1× TAE) for 25–30 min. Inspect gel as usual for indications of degradation and/or DNA contamination.

Note 1: EtBr is a mutagen that must be handled with care and disposed of properly (Appendix E).

Note 2: When working with a transilluminator or hand-held UV light source, be sure to protect eyes and skin from ultraviolet light.

7. When performing quality control checks on RNA preps, photodocumentation (Chapter 10) is extremely important for future reference.

Quality control technique 3: UV shadowing

On occasion, the amount or concentration of a nucleic sample may be too low to observe by electrophoresis, even with the very sensitive SYBR family of dyes. Regardless of the amount of the sample, it is still important to ascertain the integrity of the sample. An alternative visualization method, known as UV shadowing, is perfect for observing very low amounts on sample on a gel, particularly low molecular weight molecules run on a polyacrylamide gel. All that

⁶50× TAE (per liter) = 242 g Tris base; 100 ml 0.5M Na₂-EDTA, pH 8.0; 57.1 ml glacial acetic acid; autoclave. Melt agarose in 1× TAE buffer, and also use as the electrophoresis running buffer at a 1× concentration.

is required to UV shadow a gel is a hand-held UV light source (254 nm), a thin layer chromatography (TLC) fluor-coated silica gel sheet, and proper safety equipment (UV glasses and gloves).

Protocol

1. Carefully remove the gel from the glass plates used for the electrophoresis and place it on a sheet of plastic wrap. If necessary, gently stretch the plastic wrap to remove wrinkles that may have formed under the gel.

2. Carefully position the plastic wrap (and gel) on top of a fluor-coated TLC plate.

Note: Think about designating a single TLC plate for UV shadowing and using it exclusively for that purpose. This is an economical way to perform UV shadowing routinely, and keeping the plate near the image analysis station will also save time.

3. Taking care to avoid skin and eye exposure to UV light, position the hand-held UV monitor above the gel.

Note 1: It is helpful to place the UV light source on a slight angle above the gel in order to see the bands.

Note 2: Rather than holding the UV light source, consider clamping it onto a ring stand so that both hands remain free.

4. Mark the location of the band(s) of interest if they are to be cut out of the gel for further purification. This is easily accomplished by placing a second sheet of plastic wrap on top of the gel and then marking the position of band(s) with a lab marker.

Note 1: RNA and DNA bands appear as very dark blue bands against a pale green fluorescent background.

Note 2: Be sure to note the position of the loading dyes bromophenol blue and/or xylene cyanol, which are often mistaken for nucleic acid bands. Consider running a lane with only the loading dye(s) for comparison with the other samples.

5. Photograph the gel with the position of the loading dyes clearly indicated. This will preclude confusion when referring back to the photograph at later date.

Two common applications come to mind when thinking about UV shadowing. First, synthesis of a sense (+) or antisense (–) RNA by *in vitro* transcription is always followed by electrophoresis in order to demonstrate that the transcript is intact and of the predicted size. UV shadowing, because of its sensitivity, requires that very little sample be run on a gel, and the shadowing procedure will also display contaminating truncated transcription products, even if they are of very low abundance. Second, many laboratories now prepare their own primers to support various PCR and RT-PCR applications. It is wise to electrophorese and examine an aliquot of each new primer prior to use to demonstrate successful synthesis and to look for the presence of non-full-length oligonucleotides, commonly known as failure sequences (Fig. 6.6). Depending on the application, failure sequences are readily ousted from among the synthesis reaction products by cartridge purification (see Chapter 18).

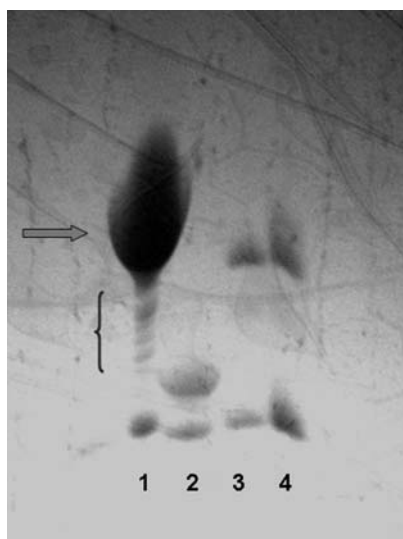


Figure 6.6 Assessment of oligonucleotide purity by UV shadowing. The major product (22-mer) in lane 1 is indicated by the arrow. Note the presence of the failure sequences (bracket) directly beneath the major band. This oligonucleotide was only desalted, which does nothing to remove non-full-length molecules. None of these bands were visible upon staining with SYBR Green even though the lanes were intentionally overloaded. Material visible in lanes 2–4 is unrelated to the quality control check. Photograph courtesy of Aptagen, LLC (Jacobus, PA).

Quality control technique 4: Sample capacity to support RT-PCR

The polymerase chain reaction (PCR) has become a mainstream tool for the molecular biologist, and its use is widespread. Without a doubt, it has revolutionized the way most problems pertaining to cell biology are approached. As such, it is likely that in any given laboratory, PCR is being used to amplify one or another transcript⁷, using primers and conditions that have long since been optimized. The frustration index rises when one is unable to generate a product using a freshly prepared sample of RNA, or an RNA sample that “used to work”. As such, one standard protocol in this laboratory is to use primers that are known to work all the time when RNA is isolated intact and clean, from a particular biological source. Testing a sample in this way can be performed using very boring primers, such as those for β -actin or GAPDH, just to show that one can at least amplify *something* from the RNA in question. If a PCR

⁷Techniques pertaining to RNA-based PCR (RT-PCR) applications are described in detail in Chapters 17–19.

product is evident with control primers, then the reason for failure to generate a product from experimental primers can be investigated more intelligently. In summary: no signal from a fairly abundant template probably means that the RNA sample is poor and unable to support reverse transcription or amplification of anything. This latter possibility will have been validated, or ruled out, through the use of these “control” primers.

Quality control technique 5: Northern analysis

Northern analysis is a most useful experimental tool by which to obtain both a qualitative and a semi-quantitative profile of any RNA sample. In the context of using Northern analysis as a quality control device, the stringent hybridization of a labeled oligo(dT)- or poly(T) probe to total RNA or poly(A)⁺ purified material will produce an image, by autoradiography or by chemiluminescence, which describes the size distribution of the polyadenylated component. Subsequently, the poly(T) probe can be removed from the blot by high-stringency washing so that the same blot can be used again for hybridization with gene-specific probes. Alternatively, aliquots of purified RNA can be probed with labeled poly(T) *before* electrophoresis to normalize RNA samples with respect to poly(A) content when the mass of the RNA is so small that it precludes normalization by conventional methodologies, such as affinity selection (Chapter 5). A detailed protocol for using a labeled poly(T) probe to normalize samples prior to electrophoresis is presented in Chapter 9. Of course, one may simply use a probe corresponding to a specific gene of interest, rather than assessing overall poly(A) content. The idea behind QC is simply to demonstrate that the sample will function in various downstream applications.

Quality control technique 6: Sample capacity to support *in vitro* translation

The biological integrity of mRNA is reflected by its ability to direct the translation of its encoded polypeptide. This ability may be assayed either by microinjection into a suitable cell type (*Xenopus* oocytes, for example) or in a cell-free translation system, such as a reticulocyte lysate or wheat germ extract. Biologically competent mRNAs will, when translated, generate a defined and reproducible pattern of protein bands. This technique, commonly known as *in vitro* translation is often coupled with immunoprecipitation, predicated upon the availability of a suitable antibody. Whereas this technique is valuable as a quality control method in certain specialized applications, it is the most technically challenging and infrequently used technique. For the most part, it is not an essential quality control technique for most routine applications.

References

- Glasel, J.A. (1995). Validity of nucleic acid monitored by 260 nm/280 nm absorbance ratios. *BioTechniques* **18**, 62–63.
- Huberman, J.A. (1995). Importance of measuring nucleic acid absorbance at 240 nm as well as at 260 and 280 nm. *BioTechniques* **18**, 636.
- Manchester, K.L. (1995). Value of A₂₆₀/A₂₈₀ ratios for measurement of purity of nucleic acids. *BioTechniques* **19**, 208–210.
- Manchester, K.L. (1996). Use of UV methods for the measurement of protein and nucleic acid concentrations. *BioTechniques* **20**, 968–970.
- Wienand, U., Z. Schwarz, and G. Felix. (1978). Electrophoretic elution of nucleic acids from gels adapted for subsequent biological tests: Application for analysis of mRNAs from maize endosperm. *FEBS Lett.* **98**, 319–323.
- Wilfinger, W.W., K. Mackey, and P. Chomczynski. (1997). Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques* **22**, 474–481.

7 Resilient Ribonucleases

Rationale	155	Other reagents used to control nuclease activity	165
Elimination of ribonuclease activity	156	Guanidine hydrochloride	165
Latent RNase contamination issues	157	Guanidine thiocyanate	165
Types of ribonuclease inhibitors	158	Sodium dodecyl sulfate	166
Specific inhibitors	158	<i>N</i> -Laurylsarcosine	166
Vanadyl ribonucleoside complexes (VDR; VRC)	158	Phenol:Chloroform: isoamyl alcohol	166
RNasin®	159	8-Hydroxyquinoline	167
Non-specific inhibitors	160	Cesium chloride	167
Preparation of equipment and reagents	161	Cesium trifluoroacetate	168
Diethyl pyrocarbonate (DEPC)	162	Proteinase K	168
Alternative: sterile water	164	RNAlater®	169
Hydrogen peroxide	164	Protocol: synthesis of VDR	169
NaOH and SDS	165	References	170

Rationale

The difficulties associated with the isolation of full-length, intrinsically labile RNA are further compounded by the ubiquity of ribonuclease (RNase) activity. RNases are a family of enzymes, present in virtually all living cells, that can degrade RNA molecules through both endonucleolytic and exonucleolytic activity. This is good and bad: RNases are bad when an investigator is trying to isolate intact RNA for various applications, while RNases are quite useful in the preparation of RNA-free DNA preps and for eliminating residual RNA molecules (target and probe) in the RNase Protection Assay (Chapter 15). RNases have minimal cofactor requirements and are active over a wide pH range. These small, remarkably stable enzymes maintain their tertiary configurations by virtue of four disulfide bridges (reviewed by Blackburn and Moore, 1982) that allow these enzymes to renature following treatment with many denaturants (Sela *et al.*, 1957), and even after boiling. Particularly resilient are RNases of the pancreatic variety (RNase A-like enzymes). The fact of the matter is that the investigator has seconds, literally, to inhibit RNase either by flash freezing fresh tissue or else homogenizing it in chaotropic lysis buffer; this is likewise true for many other tissues. Clearly, it is incumbent upon the investigator to

ensure that both equipment and reagents are purged of nucleases from the onset of the experiment. For most RNA-minded molecular biologists, saying that a reagent or apparatus is sterile is more than likely a statement that it is RNase sterile, that is, RNase-free.

Many investigators often worry less about RNase degradation of their samples than in days past. This is mainly due to the ready availability of reagents that favor expedient isolation of RNA from biological sources of all types. As such, the investment of time required by many of the traditional techniques for the control of nuclease activity is often minimized and RNase is often not a concern of central importance. This is a very poor attitude. Many of the kits and other products used for RNA isolation assume that the user has had appropriate instruction in the art of creating and maintaining an RNase-free environment. Thus, the investigator needs to make a conscious effort at all times to prevent the introduction of RNase activity and, with the same zeal, suppress endogenous (cell and tissue) RNase.

Elimination of ribonuclease activity

The method selected for controlling RNase activity must, first and foremost, demonstrate compatibility with the cell lysis procedure. Nuclease inhibitors are most often added to relatively gentle lysis buffers when subcellular organelles (especially nuclei) must be purified intact. For example, one objective of a study might be the examination of nuclear RNA independently of cytoplasmic RNA, perhaps to assess the level of gene regulation. The method of nuclease inhibition must support the integrity of the RNA throughout the subsequent fractionation or purification steps, some of which can be quite time-consuming. In addition, the reagents used to inhibit RNase activity must be easily removed from purified RNA preparations so as not to interfere with subsequent manipulations. Thus, the control of RNase activity must be considered from two perspectives:

1. Extrinsic RNase activity must be controlled. Extrinsic or external sources of potential RNase contamination must be identified and neutralized from the onset of the experiment and all reagents maintained RNase-free at all times. Extrinsic or external sources include, but are not limited to, bottles and containers in which chemicals are packaged, non-nuclease-free water, gel boxes and combs, and microbial and fungal contamination of buffers. The single greatest contamination source, however, is the oil from the fingertips, which is rich in RNase activity. Consequently, rule number one when working with RNA is to WEAR GLOVES!
2. Intrinsic RNase activity must also be controlled. Beyond the potential for accidental contamination of an RNA preparation with RNases from the lab environment, one must be acutely aware of the fact that intracellular RNases, normally sequestered within the cell, are liberated upon cellular lysis. Unless they are inhibited without delay, RNases are free to initiate degradation of the RNA that the investigator is attempting to isolate. As endogenous RNase activity varies tremendously from one biological source to the next, the degree to which action must be taken to inhibit nuclease activity is a direct function of cell or tissue type. Knowledge of the extent

of intrinsic nuclease activity is derived from two primary sources: the salient literature and personal experience.

In all cases, actions taken to control nuclease activity need to be aggressive. Failure to do so is likely to yield a useless sample of degraded RNA. Laboratories isolating RNA with the highest degree of fidelity generally maintain a “private stock” of RNA reagents and supplies that are not used for any purpose other than RNA recovery and subsequent manipulation. Such materials should definitely not be in general circulation in the laboratory.

While no one would question the RNase inhibitory properties of many reagents used to purify RNA, of even greater concern is the stability of the RNA once the isolation reagent has been removed. For example, purified RNA is never more susceptible to nuclease attack than when stored in aqueous buffer, such as TE (10 mM Tris, 0.1 mM EDTA, pH 7.5) or water. Unfortunately, losing a valuable sample is often a pivotal lesson on the need to maintain the stability of the RNA before, during, and after its isolation from the cell. The correct way to store purified RNA is described in Chapter 2.

Latent RNase contamination issues

The term “latent RNase” refers to RNase that is (transiently) inactivated because it is complexed with an RNase inhibitor. The reason that latent RNase is a concern is because upon removal, inactivation, or dissociation of the inhibitor, the RNase can become active and proceed to decimate the RNA. This commonly occurs when an RNA-RNase complex is exposed to elevated temperatures (usually $>65^{\circ}$), strongly denaturing conditions (the presence of one of the guanidinium salts ($>3\text{M}$) or urea ($>6\text{M}$), upon exposure to pH extremes (pH <5.5 or pH >9.0), or upon oxidation. The different types of inhibitors described below each have a unique pH, salt, and temperature range over which they are function, and it is wise to heed those parameters. Many types of RNase have been identified thus far, each of which having a unique sequence specificity and degradation pattern. Moreover, the presence of a 5' cap is no guarantee of RNA stability, particularly in the presence of the RNase “cocktail” found on the surface of the skin.

A standard test for latent RNase contamination is outlined in Fig 7.1. In one tube, RNA known to be of high quality (test RNA) is dissolved in water or other non-denaturing buffer (TE, for example) and then incubated for 1 h at 37° . In a second tube, heat-inactivated (15 min at 70°) RNase inhibitor is mixed with test RNA for 1 h at 37° . At the conclusion of the incubation period the contents of both tubes are electrophoresed under denaturing conditions to look for signs of degradation (Chapters 6 and 9). This incubation can be extended to 2–16 h (overnight) if very low levels of RNase are suspected. Finally, some RNase-inhibitors can be intentionally inactivated by incubation with oxidized glutathione or by withholding DTT.

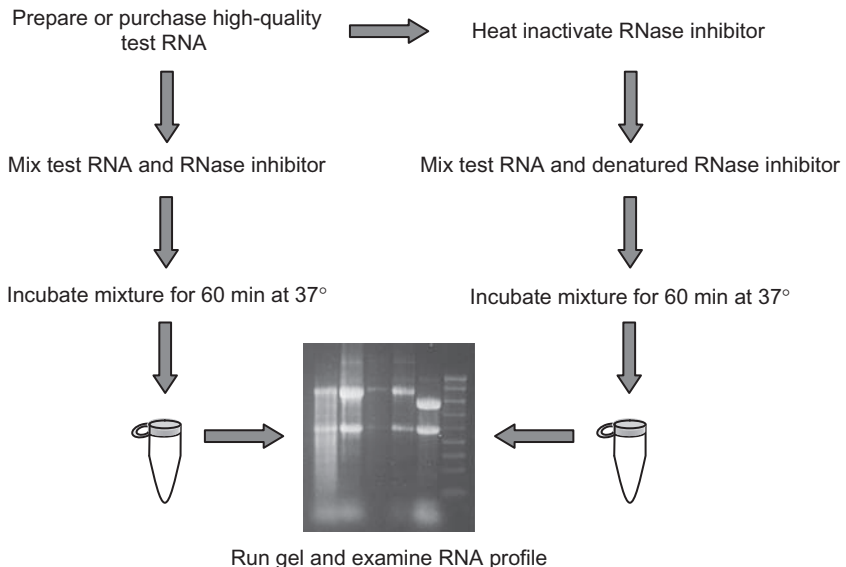


Figure 7.1 Detection of latent RNase activity. Inactivation of RNase-inhibitors will release activated RNase that may compromise the integrity of a sample. See text for details.

Types of ribonuclease inhibitors

Compounds used for control of RNase activity fall into two broad categories, specific and non-specific RNase inhibitors. Some of these compounds are added directly to lysis- and reaction buffers to control RNase activity, whereas others are used in the standard preparation of reagents intended for RNA work. The effectiveness of an RNase inhibitor in a particular system can be ascertained by a short incubation of RNase A, test RNA, and the putative inhibitor. The post-incubation products may then be viewed on a simple agarose gel (Chapter 6, Fig. 6.5). If the RNA is degraded, a change in protocol is clearly indicated. The biotech industry has fueled the demand for both broad-spectrum and application-specific nuclease inhibitors and many such products are widely available.

Specific inhibitors

Vanadyl ribonucleoside complexes (VDR; VRC)

Vanadyl ribonucleoside complexes (VDR) were developed in the mid-1970s as a means of controlling RNase activity when using relatively gentle method of cellular lysis. VDR is not widely used any longer for three reasons: (1) widespread use of harsh lysis reagents that are efficient inhibitors of RNase; (2) VDR in trace quantities can inhibit *in vitro* translation of purified mRNA (Berger and Birkenmier, 1979; Berger *et al.*, 1980); (3) VDR can interfere with cDNA synthesis by competing with dNTPs when present at low concentrations.

Therefore, VDR is contraindicated for RT-PCR and the numerous permutations thereof, and its use is not recommended.

VDR consists of complexes formed between the oxovanadium ion and any or all of the four ribonucleosides (Leinhard *et al.*, 1971; Berger and Birkenmier, 1979) in which vanadium takes the place of phosphate. These complexes then function as transition-state analogs. In the absence of VDR, RNase-mediated cleavage of the phosphodiester backbone of RNA results in the transient formation of a dicyclic transition-state intermediate, which is subsequently opened up by reaction with a water molecule. In its capacity as an RNA analog, the vanadium complex forms a highly stable dicyclic species to which the enzyme remains irreversibly bound. Thus, VDR inhibits nuclease activity by locking the RNase and its "pseudo-substrate" in the transition state. VDR binds tightly to a broad spectrum of cellular RNases and is compatible with a variety of cell fractionation methods, including the older sucrose gradient centrifugation approach (Benecke *et al.*, 1978; Nevins and Darnell, 1978). Succinctly, VDR inhibits RNase by binding to it irreversibly. VDR has been used in this laboratory with gentle NP-40 lysis buffers, though we have found that keeping all reagents ice-cold throughout the isolation procedure is equally as effective. While the habitual use of VDR is one of the more cost-effective means of controlling nuclease activity, in the world of PCR it is not an appropriate choice.

VDR is routinely prepared as a 200 mM stock solution (see protocol at the end of this chapter) and used at a final concentration of 5–20 mM, depending on the anticipated degree of RNase activity. It has been shown to suppress RNase activity in cells with significant levels of endogenous RNase (Berger and Birkenmier, 1979). VDR is particularly effective against pancreatic RNase A and RNase T₁, but not against RNase H. It is also inhibitory against the family of enzymes commonly known as the poly(A) polymerases, which may or may not be significant in the context of the types of experiments being undertaken. After it has neutralized RNase activity, VDR can be removed from an aqueous lysate by repeated, systematic extraction with organic extraction buffer containing 0.1% 8-hydroxyquinoline (w/v), which is itself a partial inhibitor of RNase. Alternatively, VDR can be removed by the addition of 10 equivalents of EDTA. The investigator is cautioned, however, that 1 equivalent of EDTA is more than sufficient to dissociate the VDR complex. Thus, when VDR is present, EDTA-containing buffers are to be avoided until another method of controlling RNase activity has been employed, for example, the addition of phenol:chloroform to the lysate. Moreover, excess EDTA carrying over can inhibit PCR dramatically (discussed in Chapters 18 and 19).

RNasin[®]

An excellent alternative to VDR are the commercially available peptides such as RNasin[®] (Promega, Madison, WI) that can be used to eliminate nuclease activity and circumvent many of the problems commonly associated with the use of VDR. Originally purified from human placenta and rat liver, the cloning

of this gene and similar genes from other sources has ensured an ample supply of these very effective peptides. RNasin is a 51,000 MW dithiothreitol (DTT)-dependent¹ protein. RNasin inactivates RNase by competitive, non-covalent binding to the enzyme with a binding constant $K_i = 4.4 \times 10^{14}$ (Blackburn, 1979; Blackburn *et al.*, 1977). Non-recombinant peptides purified from hog liver may still be purchased. Because the porcine peptide offers no risk of human DNA contamination, it may be suitable for some clinical and diagnostic research involving RNA from *Homo sapiens*.

RNasin is routinely used at a concentration of 250–1000 U/ml and is effective against RNase A, RNase B, and RNase C; it does not inhibit RNase T₁, S1 nuclease, or RNase from *Aspergillus* (Promega product insert). Newer formulations of non-specific nuclease inhibitors are also available. For example, SUPERase•In™ (Ambion, Austin TX) is a non-DTT-dependent formulation which offers broader protection against RNase. In this regard, the common generic abbreviations in the literature are RI (ribonuclease inhibitor), RIP (ribonuclease inhibitor protein) and hPRI (human placental ribonuclease inhibitor).

A major advantage of this approach to controlling nuclease activity is its compatibility with a greater variety of *in vitro* reactions, compared to VDR. These reactions include protection of template messenger RNA (mRNA) during the synthesis of cDNA (de Martynoff *et al.*, 1980) for all contemporary applications, *in vitro* transcription, in which the yield of RNA and its integrity are maintained (Scheider *et al.*, 1988), and *in vitro* translation, in which the yield of large polypeptides is improved, as is the activity of purified polysomes (Scheele and Blackburn, 1979). As noted above, care must be taken to avoid powerful denaturing conditions, such as high concentrations of urea or heating to 65°, which are known to cause the uncoupling of RNase–RNasin complexes and accompanying restoration of RNase degradative activity.

Non-specific inhibitors

The use of a variety of different types of non-specific inhibitors has been widely reported. These compounds purportedly function by eliminating all intrinsic enzymatic activity and have been incorporated into RNA isolation procedures with varying degrees of success. Although none of these has been found to be completely satisfactory for every application, one or more of these compounds may provide a means of controlling RNase activity that would otherwise not be possible. Many of these are older and no longer in widespread use, though their persistence in the literature mandates mentioning here. These non-specific inhibitors include heparin (Palmiter *et al.*, 1970; Cox, 1979), iodoacetate, polyvinyl (dextran) sulfate, cationic surfactant (Dahle and Macfarlane, 1993), and the clays macaloid and bentonite (reviewed by Blumberg, 1987). Interestingly, in addition to its biotech and pharmaceutical applications, the use of bentonite

¹In excessive quantities, DTT can be detrimental in some downstream applications. Be sure to check for DTT-compatibility with ensuing protocols.

in winemaking (Blade and Boulton, 1988) is nearly universal, dating back to the mid-1930s (Saywell, 1934). The preparation of these clays for the elimination of RNase activity has been described elsewhere (Fraenkel-Conrat, *et al.*, 1961; Poulson, 1973; Blumberg, 1987) and not included here, as they are no longer widely used to control RNase.

Perhaps closer to home is the ready availability of dilute solutions (3%) of hydrogen peroxide (H_2O_2), and mixtures of NaOH (0.1N) and sodium dodecyl sulfate (SDS; 0.1–0.5%). For those investigators who do not wish to prepare RNase remedies in-house, a number of commercial preparations also await your purchase order number. These proprietary formulations include ELIMINase[®], RNase Away[®], RNaseZap[®], and RNase-Off[™], and provide very satisfactory removal of RNase from laboratory surfaces, including bench tops.

Preparation of equipment and reagents

As delineated above, rule number one when working with RNA is to wear gloves during the preparation of reagents and equipment and during the actual RNA extraction procedure. Finger greases, especially those of overly zealous lab technicians, are notoriously rich in RNase and are generally accepted as the single greatest source of RNase contamination. Further, one should not hesitate to change gloves several times during the course of an RNA-related experiment, since door knobs, micropipettors, refrigerator door handles, telephone receivers, iPods, and computer keyboards are also potential sources of nuclease contamination. It is the strict policy of this investigator to remove his gloves when stepping away from the bench and don a fresh pair when returning to the bench and the RNA protocol.

With respect to laboratory plasticware, individually wrapped serological pipettes are always preferred for RNA work as no special pre-treatment is required. Conical 15 ml and 50 ml tubes, both polypropylene and polystyrene, are considered sterile if already capped and racked by the manufacturer. Plasticware that is certified as tissue culture sterile is not likely to be a source of nuclease activity unless subsequently contaminated (be certain to wear gloves when handling these tubes for RNA work). Bulk-packed polypropylene products are potential sources of nuclease contamination, mainly due to handling and distribution from a single bag. This pertains to microfuge tubes and polypropylene micropipette tips, because they can become contaminated and, in turn, contaminate stock solutions. Thus, any non-sterile plastic product that comes into contact with an RNA sample at any time, either directly or indirectly, and which can withstand autoclaving, should be so treated. Bulk packages of microfuge tubes are best distributed with gloved hand into some type of autoclavable receptacle. Alternative, place the tubes into several 500 ml glass beakers and cover with heavy duty aluminum foil. Then, autoclave microfuge tubes for no more than 15–20 min. These microfuge tubes are then handled only with gloves and set aside exclusively for RNA work.

When glassware must be used, as may be required for the manipulation of organic (phenol and chloroform) extraction buffers, individually wrapped borosilicate glass pipettes are strongly preferred. Glassware that must be reused most certainly should be reserved exclusively for RNA work and should not be in general circulation in the laboratory. Contrary to popular belief, however, the temperature and pressure generated during the autoclaving cycle are usually not sufficient to eliminate all RNase activity. Fortunately, RNases can be destroyed quite effectively by baking in a dry-heat oven. Glassware should be cleaned scrupulously, rinsed with RNase-free water (discussed below), and then baked for 2–3 h at approximately 200°. This is a very effective method for purging glassware of RNase activity. It is important to note that not all laboratory implements can withstand the heat generated in a dry-heat oven. If there is a question about a particular type of plastic or other material, be certain to check with the manufacturer (technical assistance toll-free numbers can usually provide the necessary information) before baking. Finally, all stock- and working solutions in the lab should be labeled with a pre-determined expiration date, in addition to the date of preparation. A sad consequence of the use of out-of-date and even worse, ancient solutions, in the lab is the unintentional introduction of RNase from microorganisms that may have taken up residence in these bottles. Keep the QA/QC people happy: pay attention to expiration dates. Finally, it is prudent to maintain separate containers of chemicals as well as stock solutions for exclusive use as RNA reagents. This is most easily accomplished by aliquoting sterile reagents into suitable volumes, rather than drawing repeatedly from the stock bottles. Aliquots used once should be discarded. Although such actions may at first seem excessive, they may well preclude the accidental introduction of RNases and facilitate recovery of the highest possible quality RNA. Finally, don't forget that chemical solids should be weighed with an RNase-free spatula.

Diethyl pyrocarbonate (DEPC)

DEPC is a potent, non-specific chemical inhibitor of RNase² and its most common application has been to make water nuclease-free. **DEPC is also carcinogenic and should be handled with care and with strict attention to the safety recommendations of the manufacturer.** There remains a fair amount of resistance to using DEPC in the laboratory. It is the opinion of this Author that there are three legitimate reasons for this:

1. DEPC is toxic and not appropriate in a lab where it cannot be handled properly. This objection is with merit, as improper handling may present a serious health risk.
2. DEPC is too difficult to remove from reagents and interferes with PCR and other reactions. This is often a direct result of adding too much DEPC: in various consulting

²Because of the efficient inhibition of certain classes of proteins by histidine and tyrosine modification, DEPC was used years ago to facilitate the destruction of proteins in some of the original plasmid isolation procedures. This is an extremely hazardous technique that is to be avoided entirely.

sorties over the years laboratory personnel have confessed to this Author to using DEPC at final concentrations ranging from 0.5–5%!

3. There are no measurable problems with RNase activity at present. Such a situation is fortunate indeed, though not widespread.

Most laboratories no longer use DEPC in-house because of the hearty array of pre-made nuclease-free solutions, including water, which can be purchased from one's favorite vendor and represent a very worthwhile investment. Still, it is worth discussing the attributes and shortcomings of DEPC here in the event that a compelling reason exists for bringing this chemical into the lab.

DEPC has been used to purge reagents of nuclease activity prior to use in an experiment involving the purification of RNA because autoclaving does not ensure the complete elimination of RNase activity; buffers treated solely by autoclaving have the potential to maintain residual RNase activity. Lack of attention to this detail may very well pose a severe impediment to the generation of data that are both representative and reproducible. DEPC was also used in the past in the preparation of non-disposable glassware prior to use for RNA work; much easier and far less toxic methods are currently available (described below).

If the decision is made to use DEPC, it should be added to a final concentration of 0.05% (v/v)³. Buffers are then shaken for several hours on an orbital platform (e.g., Lab-Line Instruments) or stirred vigorously with a magnetic stirrer for 20–30 min, after which DEPC must be destroyed completely. This is most commonly accomplished by autoclaving. At autoclaving temperature and pressure, DEPC degrades into carbon dioxide and ethanol, both of which are quite volatile under these conditions. Alternative approaches include (1) maintaining certain solutions at 60° overnight⁴, a process that should be carried out in a chemical fume hood with the bottle cap slightly loosened; and (2) removing DEPC from large volumes of DEPC-treated water by boiling it for 1 h in a fume hood. Not only is DEPC carcinogenic, which alone would justify its complete removal and careful handling, but it also has a strong affinity for adenosine nucleotides. Even trace amounts of residual DEPC will result in chemical modification of the base adenine (Henderson *et al.*, 1973), thereby changing the physical properties of RNA and compromising its utility for *in vitro* translation (Ehrenberg *et al.*, 1974) and other applications, including standard blot analysis and PCR (personal observations). This is also the reason that it has never been recommended that DEPC be added directly to cell suspensions or lysates containing RNAs to be purified. Complete removal of DEPC can be further promoted by rapidly stirring the hot solutions with a nuclease-free magnetic stir bar. Frequently, if the autoclaving time was not

³Whereas previous reports called for the use of DEPC at 0.1%, reduction to 0.05% works just as well and is much easier to purge from the solution by autoclaving.

⁴Important: Do not autoclave or otherwise heat solutions that would not ordinarily be so treated, such as those containing SDS, NP-40 (Igepal CA-630), or NaOH. These non-autoclavable components may be added to complete the solution formulation *after* the other components have been autoclaved, as needed, or otherwise made nuclease-free.

adequate, one notices the distinctive odor of residual DEPC. In days gone by in this laboratory, solutions were routinely kept in the warm autoclave chamber for approximately 30 min after completion of the autoclaving cycle in order to purge them of DEPC.

Important: Assuming DEPC is to be used in the lab:

1. Do not add DEPC to any buffer containing mercaptans or 1° amine groups, with which DEPC is reactive (Berger, 1975). Perhaps the most common buffers to which DEPC exposure is to be avoided are the Tris (*tris*(hydroxymethyl)aminomethane) buffers. Instead, use DEPC-treated water (add DEPC to water; autoclave to remove DEPC) to make up the Tris-containing solution and then autoclave the solution again.
2. Buffers consisting of chemicals that are known (or suspected) to have DEPC incompatibility can be twice filtered through a nitrocellulose membrane to remove RNase activity and other trace proteins.
3. The electrophoresis unit itself, including the gel comb, casting tray, and interior of the gel box can be treated to remove RNase contamination by rinsing and soaking in DEPC-treated water (*after* autoclaving the water). *Never* expose the unit to DEPC because acrylic is not resistant to DEPC. See “H₂O₂” section, below, for an alternative approach.

Alternative: sterile water

A suitable alternative to DEPC treatment of H₂O to render it nuclease-free is to simply buy nuclease-free water from one's favorite vendor. Purified RNA may be rehydrated in this water and stored at -80°. This water is also excellent for making dilutions of nuclease-free stock solutions. Although the habitual use of purchase of nuclease-free water is generally expensive, it is readily available and an excellent, safe alternative to the use of DEPC. For those labs affiliated with hospitals it is interesting to note that water labeled “sterile water for irrigation”, is free of contaminants and is good for RNA work!

Hydrogen peroxide

The need to eliminate RNase from equipment and supplies in the lab, including gel combs, casting trays, electrophoresis chambers, and graduated cylinders is unquestioned. Soaking such implements in a 3% solution of hydrogen peroxide (H₂O₂), commonly available in pharmacies and similar stores, is a very effective and inexpensive measure. H₂O₂ is a powerful oxidizing agent that can render a surface nuclease-free by soaking for 20–30 min, followed by rinsing with copious amounts of water that, at the very least, has been autoclaved. *Do not* use the more concentrated forms of H₂O₂ (e.g. 30%) commonly available from standard chemical supply companies. At these higher concentrations H₂O₂ is extremely dangerous and may well cause irreparable damage to acrylic gel box components and other equipment, and tissue damage to the investigator. Also, be sure to avoid old solutions of H₂O₂, as they may no longer be solutions of H₂O₂!

NaOH and SDS

An older method for the removal of RNase activity from laboratory surfaces is to soak resistant laboratory implements, or wipe laboratory surfaces, with a solution containing 0.1N NaOH and up to 0.5% SDS. The reader is cautioned to handle such a formulation with care, and subsequent extensive washing is required to remove the alkali. Failure to do so could pose a health hazard to others working in the lab, not to mention the fact that residual NaOH could create a chemical environment favoring alkaline hydrolysis of an RNA sample! Even though thorough rinsing with water precludes this potential difficulty, it is not as cost-effective as H₂O₂, especially when NaOH and SDS are purchased premixed.

Other reagents used to control nuclease activity

The following is a list of reagents commonly employed to minimize or eliminate RNase activity. These reagents are generally not used alone, but rather are components of popular RNA isolation buffers. For cells enriched in RNase, and even those that are not, homogenization in lysis buffers consisting of guanidinium thiocyanate (GTC) or guanidinium hydrochloride (G·HCl) is widely accepted as the method of choice for consistency from sample to sample, though other excellent, highly denaturing methods are also presented in this laboratory guide. Because of their extremely chaotropic nature, these compounds are quite effective for disruption of many tissues as well as cells grown in culture, and are the key component of several commercial RNA isolation formulations. There are also disadvantages to lysing cells directly in guanidinium-containing buffers, namely the intermixing of nuclear and cytoplasmic message, as described in Chapter 2. In the absence of a strongly denaturing lysis buffer, RNase inhibitors, as described above, should be added to the buffer just prior to use.

Guanidine hydrochloride

Guanidine hydrochloride (G·HCl) is a strong ionic protein denaturant. At a working concentration of about 4–6M (in water) it is an excellent inhibitor of RNase activity during purification of nucleic acids from cells and whole tissue samples. Solutions containing guanidine HCl or guanidine thiocyanate are often referred to as chaotropic buffers because of their biologically disruptive nature. Cellular and organelle membranes are disrupted on contact and proteins tend to unfold and are solubilized in the presence of this compound.

Guanidine thiocyanate

Guanidine thiocyanate (GTC) is a stronger protein denaturant than guanidine hydrochloride and is the denaturant of choice for the preparation of RNA from

sources enriched in RNase activity, especially pancreatic tissue (Chirgwin *et al.*, 1979). It is routinely used at a working concentration of 4 M and often in conjunction with a reducing agent (e.g., β -mercaptoethanol; β -ME) and an ionic detergent such as sarcosyl (N-laurylsarcosine). Cellular and organelle membranes are disrupted on contact and proteins tend to unfold and are solubilized in the presence of this compound.

Sodium dodecyl sulfate

Sodium dodecyl sulfate (SDS), also known as lauryl sulfate, is an ionic detergent that is useful for the rapid disruption of biological membranes. It is a key component of many reagents used to purify nucleic acids because of its abilities to quickly disrupt the tissue architecture and to inhibit both RNase and deoxyribonuclease (DNase) activity. SDS is usually prepared as either a 10% or a 20% (w/v) stock solution and is used most often at a working concentration of 0.1% to 0.5%. The performance of this detergent can be affected significantly by its purity. SDS is easily precipitable in the presence of potassium salts and generally is not added to guanidinium buffers, as it has very low solubility in high-salt, chaotropic solutions.

N-Laurylsarcosine

N-Laurylsarcosine, also known as sarkosyl, is an ionic detergent, similar to SDS. Unlike SDS, however, sarkosyl exhibits excellent solubility in high-salt, chaotropic solutions and is therefore the detergent of choice in guanidinium-based lysis buffers. It acts to disrupt tissue and cellular ultrastructure, inhibit nuclease activity, and facilitate the disaggregation of proteins and nucleic acids in a cellular lysate.

Phenol:Chloroform: isoamyl alcohol

CAUTION: Phenol, chloroform, and combinations thereof are caustic, carcinogenic reagents that must be handled with extreme care. Be sure to observe all safety precautions indicated by the manufacturer.

Phenol and chloroform are organic solvents that very efficiently denature and cause the precipitation of proteins. Molecular biology grade phenol (redistilled to remove impurities)⁵ is itself a good protein denaturant, although it is quite unstable. Phenol oxidizes rapidly into quinones, which impart a pinkish

⁵Commercially redistilled phenol is widely available, precluding the need for any lab to perform in-house phenol purification. While commonplace 25 years ago, the dangers of phenol distillation far out-weigh the purchase cost. The distillation of phenol in the laboratory is a very dangerous process and is to be avoided. If for some absolutely compelling reason it must be performed, the details of in-lab phenol distillation can be found elsewhere (Wallace, 1987) and should be performed by experienced personnel only.

tint to phenol solutions; if observed, pink-tinted phenol solutions should be discarded immediately. Quinones form free radicals that break phosphodiester linkages and crosslink nucleic acids. To reduce the rate of phenol oxidation, 8-hydroxyquinoline may be added to a final concentration of 0.1% (w/v). Phenol reagents are usually prepared by saturating the phenol with aqueous buffer and then adding an equal volume of chloroform. The addition of chloroform stabilizes the phenol, imparts a greater density to this organic extracting material, improves the efficiency of deproteinization of the sample, and also facilitates removal of lipids from the lysate. Mixtures of phenol and chloroform have also been shown to increase the yield of poly(A)⁺ mRNA over phenol alone (Perry *et al.*, 1972). Very often, isoamyl alcohol is used in conjunction with chloroform or with mixtures of phenol and chloroform. Isoamyl alcohol reduces the foaming of proteins that would normally be generated by the mechanics of the extraction procedure. The most common organic extraction buffer formulations consist in part of phenol:chloroform:isoamyl alcohol in a 25:24:1 ratio. In the presence of these solvents, RNase activity is inhibited.

For a more thorough discussion of the preparation and use of phenol, see Appendix D.

8-Hydroxyquinoline

8-Hydroxyquinoline is a partial inhibitor of RNase and is occasionally added to organic extraction buffers that contain phenol (Kirby, 1956). As an antioxidant, 8-hydroxyquinoline stabilizes phenol and retards the formation of quinones (phenol oxidation products). It is usually added to a final concentration of 0.1% (w/v). 8-Hydroxyquinoline imparts a bright yellow color to the phenol:chloroform to which it is added, thereby helping the investigator keep track of the organic and aqueous phases during the nucleic acid purification process.

8-Hydroxyquinoline also chelates heavy metals, making it very useful for removing VDR from cell lysates. Upon binding VDR, the 8-hydroxyquinoline changes from yellow to dark green, and repeated extractions with 8-hydroxyquinoline-containing phenol:chloroform are often necessary. When the phenol-containing phase of the extraction buffer remains yellow, all VDR has been removed. The inclusion of 8-hydroxyquinoline in organic extracting buffers may also be advantageous even when VDR is not used because heavy metals can cause RNA degradation when they are present with RNA for extended periods. Needless to say, all reagents should have been prepared using high-purity, nuclease-free water.

Cesium chloride

Cesium chloride (CsCl), once a standard centrifugation medium for both analytical and preparative separation of nucleic acids, enjoys very limited use today. Only ultrapure nuclease-free preparations of CsCl should be purchased because CsCl has only a limited ability to inhibit RNase activity. If necessary,

impure, solid CsCl may be baked at 200° for 4–6 h to remove residual RNase activity prior to exposure to RNA. The use of CsCl as a menstruum for the purification of nucleic acids in general has declined significantly in favor of newer resins and silica filters, the use of either of which precludes the requirement for density gradient (isopycnic) centrifugation (discussed in Appendix O).

Cesium trifluoroacetate

Cesium trifluoroacetate (CsTFA) is a highly soluble salt that solubilizes and dissociates proteins from nucleic acids without the use of detergents. Thus, CsTFA (available from GE Healthcare) is an excellent inhibitor of RNase and its use precludes removal of proteins from a sample by more traditional methods such as phenol:chloroform extraction or incubation with proteinase K. CsTFA is more chaotropic than CsCl, inhibits RNase to a greater extent than CsCl (Carter *et al.*, 1983), and shows greater solubility in ethanol, which expedites its removal following isopycnic centrifugation of RNA. As desired, RNA can be banded or pelleted in CsTFA because solution densities up to 2.6 g/ml are possible using this reagent. A major disadvantage of this approach, as with CsCl, is the requirement for ultracentrifugation, a technique used infrequently in the contemporary molecular biology laboratory.

Proteinase K

Proteinase K is a proteolytic enzyme (a serine protease) that is purified from the mold *Tritirachium album*. In solution, it is stable over a pH range 4.0–12.5 with an optimum of pH 8.0, and a temperature range 25–65° (Ebeling *et al.*, 1974). Although the enzyme has two binding sites for Ca⁺⁺, in the absence of this divalent cation some catalytic activity is retained. Maximum proteinase K activity is observed with the inclusion of 1 mM Ca⁺⁺ in the reaction buffer. Proteinase K digestion is routinely performed at 50°, and occasionally in the presence of EDTA in order to inhibit labile, Mg⁺⁺-dependent nucleases. Proteinase K is prepared commonly as a 20 mg/ml stock solution in sterile water (stable for 1 year at –20°) or in a solution of 50 mM Tris, pH 8.0, 1 mM CaCl₂ (stable for months at 4°). It is generally used at a working concentration of up to 50 µg/ml in any of a number of buffer formulations, including those that contain up to 0.5% SDS.

Pronase (*Streptomyces griseus*) can be used in place of proteinase K. In some cases it may be necessary to perform a pronase self-digestion to eliminate contaminating RNase and DNase activity. If necessary, this is easily accomplished by incubation of the pronase stock (20 mg/ml in 10 mM Tris-Cl, pH 7.5; 10 mM NaCl) at 37°C for 1 h. This extra task can usually be avoided by the purchase of predigested pronase, readily available from most vendors. In either case, suitable aliquots of pronase are stored at –20°. Reaction conditions for pronase are identical to those for proteinase K except that the recommended working concentration for pronase is about 1 mg/ml.

RNAlater®

RNAlater® is a proprietary aqueous formulation (Ambion) that is well known in RNA circles for its ability to prevent chemical and RNase-mediated deterioration of RNA in cells, tissues, and organs until such time as it is convenient for the investigator to process the biological material and recover high-quality RNA for any of several uses. RNAlater rapidly permeates tissues, offering several strategic advantages with respect to the collection of samples, especially when working out in the field. RNAlater is compatible with plant and animal cells and tissues, as well as *E. coli*. RNAlater has also been shown to preserve histological organization of tissue comparable to formalin-fixed and frozen sectioned tissues (Florell *et al.*, 2001).

Protocol: synthesis of VDR

The protocol described here is a modification of the procedures described by Leinhard *et al.* (1971) and Berger and Birkenmier (1979). In days gone by in this laboratory, synthesis of VDR was performed by mixing all reagents in a three-neck flask (Fig. 7.2) submerged in a hot oil bath. The three-neck configuration



Figure 7.2 The three-neck flask configuration permits easy access to the reaction components and mounting of a pH electrode, thermometer, and N₂ stream. Photograph courtesy of Corning Glass, Corning, New York.

permits access of a thermometer and pH probe, while allowing the dropwise addition of NaOH. The volumes and masses indicated may be increased proportionately to accommodate larger scale preparations.

1. Mix 0.5 mmol each of adenosine, guanosine, cytidine, and uridine in a total volume of 8 ml H₂O.
2. Heat this mixture in a heating mantle, boiling water bath, or hot oil bath until all solids are dissolved.

CAUTION: Perform this heating under a chemical fume hood and be sure to seek the guidance of someone who is familiar with the proper handling of a hot oil bath and heating mantle. Improper or careless handling can cause serious, permanent injury.

3. While flushing or bubbling the solution with a steady stream of nitrogen gas, add 1 ml of 2M vanadium sulfate (VOSO₄).

CAUTION: Vanadium sulfate is poisonous.

4. Increase the pH of the solution to about 6.0 with the dropwise addition of 10N NaOH.

5. Increase the pH of the solution to 7.0 with the dropwise addition of 1N NaOH.

Note: Be sure to change to 1N NaOH between pH 6.0 and pH 7.0 because the pH of the solution changes very rapidly in this interval.

6. At pH 7.0, a heavy precipitate of oxovanadium (IV) is observed. Accompanying the formation of this precipitate is a radical change of color from bright blue to a characteristic green-black appearance.

Note: In the absence of a nitrogen atmosphere, rapid oxidation of the oxovanadium (IV) species to the oxovanadium (V) species occurs above pH 3.5, the latter being a less efficient inhibitor of RNase.

7. Increase the final volume of the mixture to 10 ml with H₂O.

Optional: A by-product of this synthesis reaction, Na₂SO₄, can now be removed by brief centrifugation of the newly synthesized VDR, or just prior to use.

8. The product of this synthesis reaction is a 200 mM stock solution of vanadyl ribonucleoside complexes. Store small aliquots of this preparation under N₂ at -20° for 1-2 months. At -80° these complexes are stable for at least 1 year.

References

- Benecke, B.J., A. Ben-Ze'ev, and S. Penman. (1978). The control of mRNA production, translation, and turnover in suspended and reattached anchorage-dependent fibroblasts. *Cell* **14**, 931-939.
- Berger, S.L. (1975). Diethyl pyrocarbonate: an examination of its properties in buffered solutions with a new assay technique. *Analyt. Biochem.* **67**, 428-437.
- Berger, S.L. and C.S. Birkenmier. (1979). Inhibition of intractable nucleases with ribonucleoside-vanadyl complexes: isolation of messenger ribonucleic acid from resting lymphocytes. *Biochemistry* **18**(23), 5143-5149.
- Berger, S.L., M.J. Hitchcock, K.C. Zoon, C.S. Birkenmier, R.M. Friedman, and E.H. Chang. (1980). Characterization of interferon messenger RNA synthesis in Namalva cells. *J. Biol. Chem.* **255**, 2955-2961.
- Blackburn, P. (1979). Ribonuclease inhibitor from human placenta: rapid purification assay. *J. Biol. Chem.* **254**, 12484-12487.

- Blackburn, P. and S. Moore. (1982). Pancreatic ribonuclease. In "The Enzymes" (P.D. Boyer, Ed.), 3rd edition. Academic Press, New York. pp. 317–433.
- Blackburn, P., G. Wilson, and S. Moore. (1977). Ribonuclease inhibitor from placenta. Purification and properties. *J. Biol. Chem.* **252**, 5904–5910.
- Blade, W.H. and R. Boulton. (1988). Absorption of protein by bentonite in a model wine solution. *Am. J. Vitic.* **39**(3), 193–199.
- Blumberg, D.D. (1987). Creating a ribonuclease-free environment. *Methods Enzymol.* **152**, 20–24.
- Carter, C., V.G. Britton, and L. Haff. (1983). CsTFA: a centrifugation medium for nucleic acid isolation. *BioTechniques* **1**, 142–147.
- Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299.
- Cox, R. (1979). Quantitation of elongating form A and B RNA polymerases on chick oviduct nuclei and effects of estradiol. *Cell* **7**, 455–465.
- Dahle, C.E., and D.E. Macfarlane. (1993). Isolation of RNA from cells in culture using Catrimox-14 cationic surfactant. *BioTechniques* **15**, 1102–1105.
- de Martynoff, G., E. Pays, and G. Vassart. (1980). Synthesis of a full length DNA complementary to thyroglobulin 33S mRNA. *Biochem. Biophys. Res. Comm.* **93**, 645–647.
- Ebeling, W., N. Hennrich, M. Klockow, H. Metz, H.D. Orth, and H. Lang. (1974). Proteinase K from *Tritirachium album* limber. *Eur. J. Biochem.* **47**, 91–97.
- Ehrenberg, L., I. Fedorcsak and F. Solymosy. (1974). Diethyl pyrocarbonate in nucleic acid research. *Prog. Nucleic Acid Res. Mol. Biol.* **16**, 189–262.
- Florell, S.R., C.M. Coffin, J.A. Holden, J.W. Zimmerman, J.W. Gerwels, B.K. Summers, D.A. Jones, and S.A. Leachman. (2001). Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. *Mod. Pathol.* **14**, 116–128.
- Fraenkel-Conrat, H., B. Singer, and A. Tsugita. (1961). Purification of viral RNA by means of bentonite. *Virology* **14**, 54–58.
- Henderson, R.E.L., L.H. Kirkegaard, and N.J. Leonard. (1973). Reaction of diethyl pyrocarbonate with nucleic acid components. Adenosine-containing nucleotides and dinucleoside phosphates. *Biochim. Biophys. Acta* **294**, 356–364.
- Kirby, K.S. (1956). A new method for isolation of ribonucleic acids from mammalian tissue. *Biochem. J.* **64**, 405–408.
- Leinhard, G.E., I.I. Secemski, K.A. Koehler, and R.N. Lindquist. (1971). Enzymatic catalysis and the transition state theory of reaction rates: transition state analogs. *Cold Spring Harbor Symp. Quant. Biol.* **36**, 45–51.
- Nevins, J.R. and J.E. Darnell. (1978). Steps in the processing of Ad2 mRNA: poly(A⁺) nuclear sequences are conserved and poly(A⁺) addition precedes splicing. *Cell* **15**, 1477–1493.
- Palmiter, R., A. Christensen, and R. Schimke. (1970). Organization of polysomes from pre-existing ribosomes in chick oviduct by a secondary administration of either estradiol or progesterone. *J. Biol. Chem.* **245**, 833–845.
- Perry, R.P., J. La Torre, D.F. Kelley, and J.R. Greenberg. (1972). On the lability of poly(A) sequence during extraction of messenger RNA from polyribosomes. *Biochim. Biophys. Acta* **262**, 220–226.
- Poulson, R. (1973). Isolation, purification, and fractionation of RNA. In "The Ribonucleic Acids" (P.R. Stewart and D.S. Letham, Eds). Springer-Verlag, New York, pp. 243–261.

- Saywell, L.G. (1934). Clarification of wine. *Ind. Eng. Chem.* **26**, 981–982.
- Scheele, G. and P. Blackburn. (1979). Role of mammalian RNase inhibitor in cell-free protein synthesis. *Proc. Natl. Acad. Sci.* **76**, 4898–4902.
- Scheider, R., E. Schneider-Scherzer, M. Thurnher, B. Auer, and M. Schweiger. (1988). The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses a novel, highly diversified protein superfamily with a common repetitive module. *EMBO J.* **7**, 4151–4156.
- Sela, M., C.B. Anfinsen, and W.F. Harrington. (1957). The correlation of ribonuclease activity with specific aspects of tertiary structure. *Biochim. Biophys. Acta* **26**, 502–512.
- Wallace, D.M. (1987). Large- and small scale phenol extractions. *Methods Enzymol.* **152**, 33–41.

8 Stringency: Conditions that Influence Nucleic Acid Structure

Types of double-stranded molecules	173	Effect of temperature on stringency	177
Importance of controlling stringency	174	Effect of formamide on stringency	177
Effect of salt on stringency	176	Effect of urea on stringency	177
Effect of pH on stringency	176	References	178

Types of double-stranded molecules

One of the principal tools of the molecular biologist is hybridization technology, that is, the ability to promote the formation of double-stranded molecules from two complementary single-stranded nucleic acid molecules. *In vitro* and *in vivo*, this technology allows the identification and recovery of unique sequences from a mixed population of DNA or RNA molecules and is at the very heart classical and contemporary molecular biotechnologies. In this regard, the following terms have essentially the same meaning and are often used interchangeably:

hybridization
base-pairing
duplex formation
annealing.

It is important understand that the formation of a double-stranded molecule is solely a function of the bases, and not the backbone of the molecule (polyribonucleotide or polydeoxyribonucleotide). Thus, the molecular biologist can promote hybridization between any two complementary polynucleotide strands, resulting in the following types of duplexes:

Type of molecule	Symbol
double-stranded DNA	dsDNA or DNA:DNA
DNA:RNA hybrid	DNA:RNA
double-stranded RNA	dsRNA or RNA:RNA

Thermodynamically speaking¹, double-stranded DNA molecules are the least stable of the three types. DNA:RNA hybrids are more stable than double-stranded DNA because of the involvement of the RNA 2'-OH in hydrogen bond formation and, for the same reason, the most stable hybrid forms between two complementary strands of RNA. Moreover, while RNA molecules exist primarily in single-stranded form *in vivo*, the formation of extensive intramolecular base-pairing occurs is well-documented. In an experimental context, the formation of double-stranded RNA molecules permits very stringent hybridization and wash conditions. For this reason, RNA probes are often desirable as a means of increasing the signal-to-noise ratio in an experiment. If not carefully controlled, however, RNA probe molecules can participate in extensive and highly problematic non-specific base-pairing which will produce extensive background hybridization.

Importance of controlling stringency

The individual strands of a double-stranded nucleic acid molecule are dynamic indeed. Several variables, all of which can be manipulated in the laboratory, affect the ability of two single-stranded nucleic acid molecules to form stable hydrogen bonds between their complementary bases. In the laboratory, it is essential that the investigator be able to promote or prevent the formation of a double-stranded molecule, and it is often necessary to alternate between the double-stranded and single-stranded forms of a molecule in the course of a single experiment. Molecular biologists have adopted the term “stringency” to describe a defined set of conditions that are used to control the “strandedness” of nucleic acid molecules (Fig. 8.1).

Stringency is a measure of the likelihood that a double-stranded nucleic acid molecule will dissociate into its constituent single strands. Thus, stringency is all about the behavior of complementary nucleic acid molecules under what might

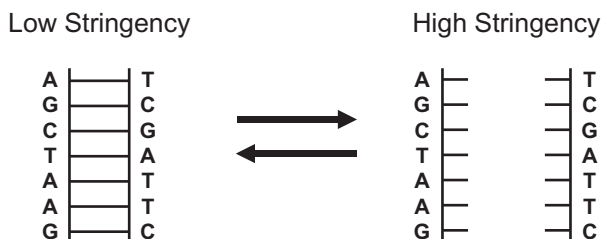


Figure 8.1 Effect of stringency on nucleic acid structure. High stringency conditions favor single-strandedness while low stringency conditions promote base-pairing between complementary molecules.

¹The stability of a duplex is also a direct function of the G + C content of the molecules involved.

be considered fairly hostile conditions. In practice, high-stringency conditions favor stable hybridization only between nucleic acid molecules with a high degree of homology, or so-called “exact-match” sequences. As the stringency in a system is lowered, a proportional increase in non-specific hybridization is favored, meaning that “mismatches” among the strands will be tolerated; mismatches are locations where a base on one molecule is not complementary to the base across from it in a double-stranded molecule, which can cause a distortion of the helix.

In vitro, the variables that are commonly manipulated to either promote or prevent nucleic acid hybridization include pH, ionic strength, temperature, and the inclusion of formamide and urea (Table 8.1). These extrinsic variables act as a function of time, meaning that the conversion from the double-stranded to the single-stranded form, and vice versa, while often rapid, is not instantaneous. The nitrogenous base composition, the length of the molecule (number of constituent nucleotides) and the degree to which two single-stranded molecules are complementary, all influence the stability of the duplex at any given level of stringency. Practically speaking, manipulating these variables is really fine-tuning the ability of nucleic acid molecules to discriminate between other molecules that have a high degree of complementarity and those that exhibit a lesser degree of complementarity.

Frequently, the temperature required to create a particular high-stringency environment is so great that the chemical stability of the RNA or DNA molecule itself would be at risk. To preclude this potentially serious difficulty, the formamide can be used to create high stringency conditions at temperatures sufficiently low so as to be unlikely to compromise the integrity of the molecules. The function of formamide is to lower the melting temperature (T_m) of base-paired molecules by interfering with hydrogen bonding. The T_m of a double-stranded molecule is best thought of as an equilibrium point, that is, that temperature at which 50% of the double-stranded molecules will dissociate into their constituent single-strands; high-stringency hybridizations are often performed only a few degrees below the calculated T_m to ensure complete hybridization wherever possible. The net result of the inclusion of formamide is a favorable reduction in the temperature needed to perform high-stringency hybridizations, thereby affording the investigator greater latitude with respect to experimental parameters.

Table 8.1 Factors that Influence Nucleic Acid Structure

Low stringency	High stringency
High salt	Low salt
Slightly acidic pH	Alkaline pH
Low temperature	High temperature

The formation of double-stranded molecules is promoted in a low-stringency environment, while dissociation into constituent single strands is favored as the stringency is increased. The addition of formamide or urea permits maintenance of high-stringency conditions at temperatures lower than would otherwise be required.

The inclusion of formamide as a means of modulating stringency can be complicated; easier methods for achieving the same end include raising or lowering the salt concentration and/or raising or lowering the temperature of the system. The take-home lesson: low-stringency conditions favor or promote double-strandedness while high-stringency conditions favor or promote single-strandedness in a mixture of nucleic acid molecules. The precise contribution of each of the stringency variables described below is discussed in much greater detail in Chapter 13.

Effect of salt on stringency

Physiological ionic strength, while somewhat variable, averages around 155 mM. In an experimental setting the natural electrostatic repulsion that would prevent two complementary, negatively-charged strands from base pairing can be neutralized with low, moderate, or high concentrations of monovalent cation, most often Na^+ , K^+ , or Li^+ . Referred to as *counterions*, the positive charge that each contributes will minimize the tendency for natural electrostatic repulsion between two negatively charged phosphodiester backbones. At low concentrations or in the absence of salt, electrostatic repulsion will easily exceed the strength of the hydrogen bonds that could otherwise form between complementary base pairs, resulting in the dissociation of double-stranded molecules, especially at elevated temperatures. In the cell, intracellular cationic proteins act, at least in part, to neutralize the negative charge associated with nucleic acids, thereby preventing inappropriate strand denaturation.

Effect of pH on stringency

pH influences the degree of ionization of nucleic acid molecules. At near neutral physiological pH, nucleic acids are ionized to an extent; hybridization can occur between complementary polynucleotides even though the individual molecules demonstrate a net negative charge. *Slightly* below physiological pH, a more acidic environment will neutralize the phosphates and cause nucleic acids, especially dsDNA, to become condensed. *In vitro*, however, a concern when establishing an acidic environment ($\text{pH} < 5.5$) is the potential that nucleic acid depurination (read:degradation) will occur in fairly short order. Above physiological pH, a more alkaline pH will increase the ionization of the phosphate groups to the extent that electrostatic repulsion can easily exceed the thermodynamic stability of a double-stranded molecule, resulting in strand dissociation. Although adding NaOH to denature DNA remains a fairly widespread practice, when working with RNA it is of paramount importance to remember the exquisite susceptibility of these molecules to degradation by alkaline hydrolysis. Do not expose RNA to a strongly alkaline environment ($\text{pH} > 8.6$) unless the intention is to destroy it.

Effect of temperature on stringency

In general, the effect of temperature is opposite that of salt concentration. By applying more heat to the system, one will eventually exceed the T_m of the duplex, causing dissociation of the strands. Moreover, the shorter the nucleic acid molecules involved, the more sensitive they are to even subtle temperature differences. Any molecular biologist can attest to the fact that oligonucleotides used as hybridization probes or primers for PCR often work (or don't work!) precisely because of what this Author refers to euphemistically as a thermodynamic miscalculation: one may easily promote or prevent hybridization by changing the temperature by as little as 1° . The thermodynamic stability of a duplex is a function not only of the number of nucleotides that participate in duplex formation but also of the guanine and cytosine (G + C) content: remember that three hydrogen bonds naturally occur between the bases guanine and cytosine and only two hydrogen bonds exist between adenine and thymine, and between adenine and uracil. The exact experimental conditions that define stringent hybridization and post-hybridization washes are addressed in Chapter 13, while the use of oligonucleotides as PCR primers is discussed in Chapters 18 and 19.

Effect of formamide on stringency

The T_m of a double-stranded molecule is largely a function of the number of base pairs involved and the G + C content, especially when the molecule is less than 100 base pairs. In cases such as these, the thermodynamic stability of the duplex is exquisitely sensitive to the temperature of the system. With longer probes, however, very high temperatures would be required to maintain a comparable level of stringency. Often, the calculated hybridization temperature (discussed in Chapter 13) is simply too high to permit base pairing under any circumstances. When this occurs, the investigator may choose to incorporate formamide into a hybridization recipe. The reason for including formamide is to lower the T_m while maintaining high stringency: a high stringency environment is created at a temperature that is low enough to permit base pairing to occur. This ensures high fidelity hybridization without the concerns associated with temperature extremes. As a general rule, one may expect to see the T_m lowered by approximately 0.75° for each 1% formamide added. While an effective additive for the T_m reduction, formamide is also a rather potent toxic compound that tends to vaporize at the elevated temperatures commonly associated with high-stringency manipulations; it should therefore be handled with great care. In some settings, the substitution of urea may be an excellent alternative.

Effect of urea on stringency

Urea has been widely used with polyacrylamide gel electrophoresis of single-stranded RNA and DNA molecules, as well as proteins. It has the desirable

qualities of being a duplex destabilizing compound (Kourilsky *et al.*, 1970) and is considerably less toxic than formamide. Thus, urea has been used successfully at concentrations ranging from 2–4M in place of 50% formamide (Simard, Lemieux, and Côté, 2001). At concentrations greater than 4M it is believed that the concomitant increase in viscosity may be responsible for diminished sensitivity of the assay (Hutton, 1977). Urea is able to participate in hydrogen bond formation and therefore is able to destabilize double-stranded polynucleotide formation by providing alternative hydrogen bonding sites for the nitrogenous bases. As a direct result of these interactions, the T_m of double-stranded molecule is reduced, making it easier to melt apart at lower temperatures. A protocol for the use of urea in agarose gel electrophoresis is presented in Chapter 9.

References

- Hutton, J.R. (1977). Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic Acids Res.* **4**, 3537–3555.
- Kourilsky, P., S. Manteuil, M.H. Zamansky and F. Gros. (1970). DNA–RNA hybridization at low temperature in the presence of urea. *Biochem. Biophys. Res. Commun.* **41**, 1080–1087.
- Simard, C., R. Lemieux and S. Côté. (2001). Urea substitutes toxic formamide as destabilizing agent in nucleic acid hybridizations with RNA probes. *Electrophoresis* **22**, 2679–2683.

9 Electrophoresis of RNA

Rationale	179	Ribosomal RNA	197
Normalization of nucleic acids	180	Gel staining techniques	204
Protocol: poly(A) normalization	183	Ethidium bromide	204
Sample preparation	183	SYBR [®] green	207
RNA denaturing systems for agarose gel electrophoresis	185	SYBR [®] gold	207
Formaldehyde denaturation	187	SYBR [®] safe	208
Protocol: formaldehyde denaturing gels	188	GelStar [®]	209
Urea denaturation	189	Silver staining	209
Protocol: urea denaturation	191	Acridine orange	210
Glyoxal/Dimethyl sulfoxide denaturation	191	Methylene blue	210
Protocol: glyoxalation and electrophoresis of RNA	192	Safety considerations and equipment maintenance	212
Running RNA on non-denaturing gels	194	Running agarose gels for the first time: a few tips	213
Molecular weight standards	194	Essential vocabulary	213
Proper use of size standards	195	Points to keep in mind	213
		References	216

Rationale

The expression and differential modulation of discrete genes and complex gene relays result in the precise orchestration of the cellular biochemistry. One fundamental approach to the study of salient transcription events involves the analysis of messenger RNA (mRNA) and precursor heterogeneous nuclear (hnRNA) as a parameter of gene expression. Such an approach is predicated on the fact that changes in the abundance of a particular RNA species may well represent a significant biochemical event in the cell cycle. At the heart of this approach is the ability of nucleic acid probes to discriminate between individual members of an extremely heterogeneous mixture of RNA molecules, and do so with high fidelity. Conditions that promote hybridization between target mRNA and a complementary nucleic acid probe may be established for the qualitative and quantitative assessment of those RNA species of immediate interest to the investigator.

Electrophoresis is a standard laboratory technique that is performed using high-voltage power supplies that have the potential to cause life-threatening injuries. Always use extreme caution when working with such equipment, and seek the assistance of someone who is familiar with safe operating procedures. Electrophoresis is a chromatography technique by which a mixture of charged molecules is separated according to size when placed in an electric field. RNA samples may be purified by any of a variety of RNA extraction procedures described in Chapters 2–5, or elsewhere. Electrophoresis is extremely versatile and is used in conjunction with a number of different assays, including Northern analysis, nuclease protection, and all forms of end-point PCR¹. For example, in the Northern analysis, colloquially known as Northern blotting, electrophoretically resolved samples of total cellular, total cytoplasmic, poly(A)⁺, and/or poly(A)⁻ RNA are transferred to (blotted), and immobilized on, the surface of a suitable membrane (Alwine *et al.*, 1977; 1979), usually nylon. The blot is then subjected to hybridization conditions in an attempt to base-pair specific RNA molecules on the filter to a complementary probe provided by the investigator. This is a variation on the theme of Southern analysis (Southern, 1975), also known as Southern blotting, in which electrophoretically separated DNA fragments are blotted onto a membrane for nucleic acid hybridization analysis.

Normalization of nucleic acids

In the preparation of RNA for electrophoresis or any type of side-by-side analysis of samples, it is incumbent upon the investigator to ensure that the samples are “normalized”, meaning that equal quantities of RNA are loaded in each lane of the gel. But what, exactly, constitutes a battery of normalized samples? Does the normalization of samples of total cellular RNA, total cytoplasmic RNA, and poly(A)⁺ mRNA have the same meaning? Is it appropriate to normalize samples by mass (equal μg quantities per lane)? What is the difference between a loading control and a transcription control? Normalization must constitute a logical and proper basis for the comparison and interpretation of hybridization signal intensity observed in each experimental sample; it must also give meaning to signals, if any, generated from positive and negative controls. The ramifications of normalization are far-reaching indeed.

There are three major issues associated with the normalization of nucleic acids.

1. Normalization based on the total mass of nucleic acid in each sample. In general, the sample with the lowest concentration is designated the “baseline” sample, meaning all of the other samples are diluted to match the concentration of the baseline sample. This is the starting point in most instances for normalization among samples.
2. Visual inspection of RNA after electrophoresis and staining. This is to ensure that each sample is intact and to ensure uniformity in the abundance of the rRNA species

¹End-point PCR refers to gel electrophoresis of PCR products after all cycles have been performed. Compare to real-time PCR. See Chapter 19 for details.

from one sample to the next, as described below. One should be aware of the fact that normalization as described above, i.e. loading equal amounts of RNA per lane, does not tell the whole story. It is important to note that rRNA is a much better loading control than a transcription control.

3. Validation of normalization by hybridization of the RNA to a “housekeeping” gene, that is, a gene whose expression is not expected to change as a function of an experimental stimulus or manipulation. Popular housekeeping genes include β -actin, fibronectin, histone, GAPDH, and transferrin receptor mRNA. Some investigators even use the total mass of tRNA as a method by which to verify normalization. The reader is cautioned, however, that there is no single transcript that is optimal for normalization because different cells respond differently to various experimental stimuli: subtle transcriptional and/or posttranscriptional variations may accompany experimentally-induced permutations in the cellular biochemistry. This aspect of the assessment of gene expression is described in detail in Chapter 12 in the section on internal controls.

The most widely accepted procedure for normalization is to measure the concentration of each RNA sample spectrophotometrically, and then load equal (A_{260}) microgram quantities of RNA in each lane on a gel. When purified poly(A)⁺ mRNA is available, loading 2 to 5 μ g per lane is sufficient to detect many low abundance mRNAs. In other cases, however, it may be desirable or even necessary to electrophorese 15 to 20 μ g of total cellular RNA or total cytoplasmic RNA without prior selection of the poly(A)⁺ RNA. For example, the yield of poly(A)⁺ RNA that can be generated in a particular model system may be so small as to preclude its isolation. Moreover, because no isolation procedure is 100% efficient, steps taken to purify poly(A)⁺ transcripts may actually serve to further under-represent low- and very low abundance mRNAs and may be contaminated with varying amounts of non-poly(A) material. Other reasons for electrophoresing total cellular RNA may include the plain and simple fact that the transcript under investigation is of sufficient abundance for easy detection, without the need for poly(A) enrichment.

The most prevalent component of a purified sample of total cellular or total cytoplasmic RNA is ribosomal RNA (rRNA), which constitutes 80 to 85% of the sample by mass. This is clearly an enormous excess compared to the mass of the mRNA transcripts, which is usually no more than 2 to 3% of the total. Thus, changes in the transcription of discrete mRNAs, even significant changes, are virtually undetectable by standard UV spectrophotometry. Yet, accurate normalization of experimental samples mandates a rationale that will ensure that any observed variations in the abundance of specific mRNAs truly indicate differential transcription or posttranscriptional regulation of these species (as opposed to an overall change in the transcription of all RNA species) and that these observed changes are meaningful. These considerations have across-the-board applicability to all assays with any quantitative component.

In order to at least partially address these concerns, one might consider the assay of poly(A)⁺ material only. Although the short-comings of this approach should be clear from the arguments presented in “The Poly(A) Caveat” (Chapter 5), loading equal amounts of poly(A)⁺ RNA (both hnRNA

and mRNA) may be considered at least one acceptable attempt to normalize samples, though this may not always be possible. When the availability of very small quantities of RNA, cells, or tissue precludes poly(A)⁺ selection, another method for normalization involves synthesis of a labeled poly(T) probe (Fornace and Mitchell, 1986; Hollander and Fornace, 1990; Farrell and Greene, 1992). The probe, prepared by performing what amounts to the synthesis of first-strand cDNA, can be used to estimate the overall poly(A) content of various samples. While not a commonly used approach, an adaptation of this high-precision technique is presented here. With this normalization approach, a minimum amount of cellular RNA (500 ng or less) is sacrificed for slot blot or dot blot analysis. The observed hybridization of the poly(T) probe is a barometer of the poly(A) content of the sample (Fig. 9.1). When compared with the hybridization signal intensity from other samples, this information may be used to normalize samples of total RNA such that equivalent quantities of poly(A)⁺ material are loaded into each lane of the gel or used for any other analysis requiring precise normalization. This technique accomplishes the equivalent of UV spectrophotometric normalization of samples of purified poly(A)⁺ mRNA; failure to normalize with respect to poly(A)⁺ content may result in observed variations that result directly from loading different quantities of poly(A)⁺ RNA into each lane.

The most important point to remember with respect to the normalization of several samples is that RNA electrophoresis and Northern analysis, while historically venerable techniques, are very close to the bottom of the sensitivity index (Chapter 19) with respect to the assay of RNA as a parameter of gene expression. Northern analysis, in which RNA electrophoresis is an initial step,

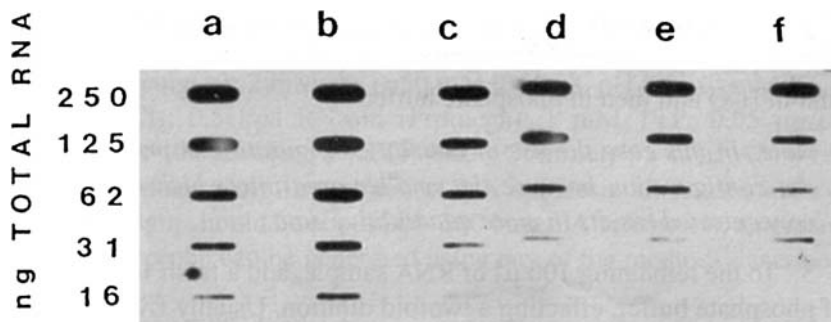


Figure 9.1 Slot blot normalization of poly(A)⁺ content in samples of total cytoplasmic RNA. Serial two-fold dilutions of 500 ng (A_{260}) were made in phosphate buffer and 100 μ l was applied to each slot. Samples were then hybridized to a ³²P-labeled poly(T) probe to determine if samples of total RNA normalized by mass contained equal mass quantities of poly(A)⁺ mRNA in the absence of oligo(dT) affinity selection. Autoradiography, coupled with scanning densitometry (not shown), revealed that although an equal mass of total RNA was applied to each slot, the poly(A)⁺ component of these samples varied from two- to four-fold.

should be used to roughly gauge the behavior of transcripts in the cell; the need for fine-tuned analysis of gene expression should then lead the investigator to super-sensitive RT-PCR-based assays.

Protocol: poly(A) normalization

Sample preparation

1. Remove 500 ng (A_{260}) of RNA from each sample to be normalized and dilute each in 200 μ l ice-cold 50 mM phosphate buffer (50 mM $Na_xH_xPO_4$), pH 6.8 prepared from the monobasic and dibasic salts (see Table 9.1 for formulation).
2. Using a Dot-Blot (Minifold I) or Slot-Blot (Minifold II) apparatus (Whatman, Inc.), apply half of each sample (100 μ l = 250 ng) to a nylon filter that was pre-wet first in nuclease-free H_2O and then in phosphate buffer.

Note: In this case the use of the slot configuration is preferred to the dot configuration because the smaller area into which the sample is concentrated results in more quantitative data.

3. To the remaining 100 μ l of RNA sample, add a fresh 100 μ l aliquot of phosphate buffer, effecting a two-fold dilution. Usually five two-fold dilutions blotted onto a filter membrane are sufficient to accurately normalize the sample.
4. Wash each well twice with 300 μ l phosphate buffer.
5. Positive controls may be generated with the dilution of commercially available pure poly(A), which, at greater dilutions, can serve as a standard for estimating the absolute poly(A)⁺ content of the sample. For an excellent negative control, prepare a similar dilution series using tRNA.
6. Immobilize the RNA on the filter paper as described by the manufacturer, ideally by UV crosslinking. Begin prehybridization or store filter in a cool, dry place until ready to use.

Prehybridization – Option 1

- 7a. Church blots (Church and Gilbert, 1984): Equilibrate filters for a minimum of 20 min in prehybridization solution (1% bovine serum albumin; 1 mM EDTA; 0.5 M $Na_xH_xPO_4$, pH 7.2, 7% SDS) at 65°. Use a minimum of 100 μ l/cm² prehybridization solution. Proceed to step 8.

Prehybridization – Option 2

- 7b. Soak filter(s) in 5 \times SSPE for at least 15 min and then prehybridize in a volume of at least 100 μ l/cm² prehybridization buffer (5 \times SSPE, 5 \times Denhardt's solution, 0.1% SDS, 100 μ g/ml denatured salmon sperm DNA, 50% formamide). Allow prehybridization to proceed for 2–3 h at 42° with gentle rocking on an orbital shaker incubator.

Synthesis of poly(T) probe

8. The poly(T) probe necessary for this experiment can be synthesized essentially as first-strand cDNA by the reverse transcription of a poly(A) template (polyadenylic acid). A typical reaction (100 μ l volume) consists of 250 ng/ μ l poly(A); 2 mM DTT; 50 mM Tris-Cl, pH 8.3; 40 mM KCl; 10 mM $MgCl_2$ 0.5 U/ μ l RNasin (Promega); 1 mM

Table 9.1 Preparation of $\text{Na}_x\text{H}_x\text{PO}_4$ Buffer

NaH_2PO_4 (ml)	Na_2HPO_4 (ml)	Resulting pH
93.5	6.5	5.7
92.0	8.0	5.8
90.0	10.0	5.9
87.7	12.3	6.0
85.0	15.0	6.1
81.5	18.5	6.2
77.5	22.5	6.3
73.5	26.5	6.4
68.5	31.5	6.5
62.5	37.5	6.6
56.5	43.5	6.7
51.0	49.0	6.8
45.0	55.0	6.9
39.0	61.0	7.0
33.0	67.0	7.1
28.0	72.0	7.2
23.0	77.0	7.3
19.0	81.0	7.4
16.0	84.0	7.5
13.0	87.0	7.6
10.5	90.5	7.7
8.5	91.5	7.8
7.0	93.0	7.9
5.3	94.7	8.0

These are the theoretical proportions of 200mM stock solutions of NaH_2PO_4 and Na_2HPO_4 required to prepare a pH-specific sodium phosphate buffer. Be certain to check pH after mixing NaH_2PO_4 and Na_2HPO_4 and subsequently diluting it. Minor adjustments in pH are frequently required. See text for details.

dTTP; 0.05 $\mu\text{g}/\mu\text{l}$ Oligo dT_{12-18} ; 100 μCi α - ^{32}P TTP (3000 Ci/mmol)² 0.5 U/ μl AMV reverse transcriptase. Any first-strand cDNA synthesis kit is satisfactory, though a higher specific activity probe will result if the cDNA synthesis reaction is performed in the presence of dTTP alone, rather than a dNTP cocktail. The probe length will be about 200–300 bases. Alternatively, a non-isotopic poly(T) probe can be generated using any of the methods described in Chapter 12.

9. Incubate at 42° for 1.5 h.

10. Terminate the reaction with the addition of 10 μl 1M NaOH and 2 μl 500mM EDTA. Incubate at 65° for 15 min to hydrolyze the poly(A) template.

Historical note: This is analogous to the removal of the RNA template following first strand synthesis during classical cDNA synthesis.

11. Neutralize the alkaline reaction mixture with 10 μl 1M Tris-Cl, pH 7.5 for 15 min at room temperature.

²Alternatively, 250 μCi of ^{35}S -thymidine 5'-(α -thio)triphosphate can be used to label the probe, though detection will not be as rapid.

12. Radiolabeled probe must be separated from unincorporated radionucleotides and other reaction components. This is easily accomplished by spun-column chromatography (Sambrook and Russell, 2001) through Sephadex™ G-25 (GE Healthcare) or by using another commercially available concentrating/purification device. Unused probe (and previously used hybridization buffer containing probe) should be stored with proper shielding at -20° , at which temperature it will be stable for 4–6 weeks.

Hybridization

13. Replace prehybridization buffer with a fresh aliquot of the same prehybridization buffer, using no more than $100\mu\text{l}/\text{cm}^2$ filter. Add the poly(T) probe at a concentration of 100 ng/ml or $1 \times 10^6\text{ cpm/ml}$.
14. Hybridize blots with gentle agitation for 3 to 4 h at 44° .

Post-hybridization washes

15. At the conclusion of the hybridization period remove the probe solution. The probe can be saved (frozen) for subsequent hybridization analyses. If filters were prehybridized and hybridized in Church blot buffer (Option 1), then wash blots twice for 10 min each in 200 mM NaCl at room temperature, once for 30 min at 44° , and finally once in 75 mM NaCl at room temperature for 5 min. If filters were hybridized in SSPE-containing buffer (Option 2), then wash the filters twice in $2 \times \text{SSPE}$, $0.1\% \text{ SDS}$ for 10 min at room temperature and then wash twice in $0.1 \times \text{SSPE}$, $0.1\% \text{ SDS}$ for 10 min at 37° . Use a volume of at least 100 ml of buffer for each wash. This will ensure efficient and effective washing to the filter(s).
16. Briefly blot the filters to remove excess wash buffer. This is readily accomplished by placing the filter, probe-side up, on a piece of Whatman 3MM paper; usually 20–30 s is adequate. Do not allow the filter to dry out.
17. Wrap the filters, while still damp, in plastic wrap and then set up autoradiography³ With this type of hybridization a single overnight exposure at -70° with an intensifying screen is quite sufficient.

Note: After an autoradiogram has been generated, each dilution can be quantified by scintillation counting: individual dilutions (i.e., the dots or slots) can be cut from the membrane and dropped directly into scintillation cocktail for measurement.

18. The poly(A) content of each sample, as determined by the degree of poly(T) probe hybridization, permits samples of total cellular or total cytoplasmic RNA to be normalized to that of a reference sample. In many cases this may simply be a control or untreated sample. All RNA samples are then prepared for electrophoresis.

RNA denaturing systems for agarose gel electrophoresis

Messenger RNA molecules *in vivo* are single-stranded polyribonucleotides complexed to cytoplasmic proteins, which together constitute messenger ribonucleoprotein (mRNP) particles. These mRNPs presumably represent the natural form of mRNA in the cell and exhibit considerable higher-level folding. This is due to RNA:protein interactions and to abundant formidable intramolecular

³Detection mechanics will be different for chemiluminescence. See Chapter 14.

base-pairing. Virtually all RNA molecules are believed to contain short double-helical regions consisting not only of the standard A::U and G::C base pairs but also weaker G::U base pairs. The extent of these intermolecular and intramolecular interactions is commonly believed to regulate the translatability of the transcripts in the cytoplasm (for review, see Lewin, 2008). Moreover, a rapidly growing body of evidence supports the notion that the transient formation of double-stranded regions of mRNA either near the 5' end of the molecule or within the body of the mRNA influence its translatability. This is the basis of antisense technology and RNA interference (RNAi). The idea is to either block interaction of the 5' end of the message with the translation apparatus or simply cleave the mRNA so that it is non-translatable.

One of the parameters that governs the migration of nucleic acids in an electric field is the molecular conformation: secondary structures or "RNA hairpins" impede electrophoretic separation based solely on molecular weight. Transfer RNA (tRNA) is the premier example of a molecule that manifests so much self-complementarity that the molecule assumes its very characteristic, easily recognized, and amazingly stable tertiary structure. Identical species of RNA molecules exhibiting varying degrees of intramolecular base pairing will migrate toward the anode (the positive electrode) at different rates, resulting in smearing rather than distinct band formation. To circumvent this problem, the electrophoresis of RNA is routinely carried out under what are known as denaturing conditions, primarily because heating alone does not solve the problem of secondary structure. The successful electrophoresis of RNA is accomplished in two parts: (1) RNA is denatured prior to loading the gel; and (2) during the electrophoresis, conditions that support and maintain the denatured state are established. These steps will ensure that like species of RNA co-migrate, and the best possible resolution is achieved when gels are poured as thin as is practical and run at low voltage.

The choice of denaturing system and gel matrix are determined primarily by the size range of the RNAs to be separated and also whether the characterization of RNA is to be preparative or strictly analytical in nature. For optimal resolution of very small nucleic acids (fewer than 500 bases or base-pairs) the matrix of choice is polyacrylamide (3–20%) (Maniatis *et al.*, 1975). Such applications routinely include the S1 nuclease assay (Berk and Sharp, 1977; Favaloro *et al.*, 1980; Sharp *et al.*, 1980), the RNase protection assay (Zinn *et al.*, 1983; Melton *et al.*, 1984) and DNA sequencing. However, for procedures such as the Northern analysis, the investigator is working with a sample containing thousands of different-size transcripts; the optimal balance between electrophoretic resolution and efficiency of transfer (blotting) from the gel to a membrane for hybridization is achieved with a 1.0 to 1.2% denaturing agarose gel.

The most commonly used RNA denaturant for agarose gel electrophoresis is formaldehyde⁴ (Boedtker, 1971; Lehrach *et al.*, 1977; Rave *et al.*, 1979). The glyoxal/dimethyl sulfoxide (DMSO) denaturing system (McMaster and

⁴Caution: formaldehyde is carcinogenic and a teratogen.

Carmichael, 1977; Bantle *et al.*, 1976; Thomas, 1980) has been used but has fallen out of favor in most labs and is therefore no longer in widespread use. As an historical note, the methylmercuric hydroxide system (Bailey and Davidson, 1976; Alwine *et al.*, 1977; 1979) was the first to be used to denature RNA for agarose electrophoresis. Ironically, while it is the most efficient RNA denaturant (Gruenwedel and Davidson, 1966), it is also by far the most toxic. Due to the great toxicity of this chemical (for toxicity review, see Junghans, 1983) it is not in use today for this application. For polyacrylamide gels, formamide⁵ (Spohr *et al.*, 1976) and urea (Reijnders *et al.*, 1973) have been used successfully as RNA denaturants. Each of these denaturing systems has unique characteristics and safety requirements as well as distinct advantages and limitations. Methodologies pertaining to the use of formaldehyde and the glyoxal/DMSO system are presented in this chapter. In addition, a novel protocol which describes the use of urea as an RNA denaturant in agarose gels is described below. In this laboratory, the urea denaturing system has been used very successfully for Northern analysis, completely eliminating the need to use formamide or formaldehyde. Because of the intrinsic difficulties and great health hazards associated with methylmercuric hydroxide it is not recommended as a denaturant and its use is not described herein.

Formaldehyde denaturation

Formaldehyde and formamide are very commonly used denaturants of RNA. Both are carcinogens and teratogens; these should be handled in a chemical fume hood and avoided altogether by expectant mothers. To ensure that the RNA migrates only with respect to molecular weight, samples of RNA are denatured with both formaldehyde and formamide prior to electrophoresis, and formaldehyde is added to the gel itself to maintain the denatured state during electrophoresis. Formaldehyde is routinely supplied as a 37% stock solution (12.3M), containing 10% to 15% methanol as a preservative. Supplied in a dark brown bottle, formaldehyde should be stored at room temperature and out of direct sunlight. Formaldehyde oxidizes when it comes into contact with the air. As such, the pH of formaldehyde should be checked prior to each use, and must be at pH greater than 4.0 for RNA work. At pH below 4.0 RNA will experience severe degradation. To raise the pH, deionize formaldehyde in the same manner as formamide (Appendix H). In many laboratory settings, formaldehyde gels furnish greater detection sensitivity than the glyoxal denaturing system, due in part to difficulties associated with the complete removal of glyoxal from the RNA prior to hybridization. The vast majority of laboratories performing RNA electrophoresis use the formaldehyde denaturing system.

The running buffer of choice for formaldehyde gels is $1 \times$ MOPS buffer. This buffer is conveniently prepared as a $10 \times$ stock (400mM MOPS⁶ pH 7.0; 100mM sodium acetate; 10mM EDTA, pH 8.0) and, if made properly,

⁵Caution: formamide is carcinogenic and a teratogen.

⁶MOPS = 3-[N-morpholino]propanesulfonic acid.

assumes a characteristic straw color when autoclaved. Contrary to popular belief, autoclaving MOPS buffer *is* acceptable and will not interfere with its role as an electrophoresis running buffer. Alternatively, MOPS buffer can be filtered twice through a nitrocellulose membrane to eliminate nuclease activity and to extend its shelf life.

Gels containing formaldehyde are less rigid than other agarose gels and are considerably more slippery. It is always wise to support such gels from beneath, as with a spatula or within the casting tray when moving them from place to place. If running an RNA gel for the first time, the information at the end of this chapter may be very useful.

Protocol: formaldehyde denaturing gels

1. Wear gloves!
2. To prepare 200 ml of molten agarose solution (more than enough for two 12 cm × 14 cm gels), melt 2.4 g of agarose in 170 ml sterile H₂O in an Erlenmeyer flask on a hot plate or in a microwave oven. Be certain that the agarose has melted *completely* by gently swirling the flask and looking for undissolved beads of agarose, commonly known as “fish-eyes.” When the gel has cooled to 55 to 60°, add a mixture of 20 ml prewarmed 10 × MOPS buffer and 10 ml prewarmed 37% formaldehyde. This will produce a solution of 1.2% agarose in 1 × MOPS buffer and 0.6 M formaldehyde. Keep this molten agarose solution in a 55–60° water bath until ready to cast the gel, and be sure to cover the flask to prevent the escape of formaldehyde fumes.

Note 1: Certain older protocols recommend a final formaldehyde concentration of 2.2 M, which is much too high. Most investigators use a maximum final concentration of 0.6 M which, if desired, may be further reduced to 0.22 M.

Note 2: In this lab, a collection of Erlenmeyer flasks are autoclaved in advance and set aside specifically for the preparation of RNA gels.

3. Cast a 0.5–0.75 cm thick gel using the solution prepared in step 2. Be sure to cast the gel in a chemical fume hood to minimize formaldehyde fumes in the room. Remember that the thinner the gel, the better the resolution of the RNA.
4. Prepare each sample for electrophoresis by mixing the following in a sterile microfuge tube:
 - 4.7 μl RNA (up to 20 μg)
 - 2.0 μl 5 × MOPS buffer
 - 3.3 μl formaldehyde
 - 10.0 μl formamide.

Be certain to mix thoroughly by pipetting and then pulse centrifuge to collect the reaction components at the bottom of the microfuge tube.

Note 1: Do not exceed 20 μg RNA per lane, beyond which one runs the risk of overloading the lanes of the gel and losing resolution. If rare or very low abundance mRNAs are of interest it may be worthwhile to load as much as 4 to 5 μg of poly(A)⁺ RNA per lane instead.

Note 2: It is imperative to use freshly deionized formaldehyde and formamide.

5. Prepare 2 to 3 μg of RNA molecular weight standard for electrophoresis, as described above in step 4.

Note: These molecular weight standards are single-stranded and need to be denatured, too.

6. Denature all samples and the molecular weight marker by heating to 55° for 15 min or at 65° for 10 min.
7. During the incubation of the RNA and once the gel has solidified, immerse the gel in running buffer (1× MOPS).

Note: Be sure to use the smallest amount of buffer necessary to completely cover the gel, including the edges which tend to “smile” by clinging to the side of the casting tray. Carefully remove the comb by slowly pulling it upward with uniform force. Covering the gel with buffer prior to removal of the comb will lubricate the wells and reduce the vacuum effect created when the comb is pulled. If the gel is not immersed in buffer, comb removal may damage the bottom of the wells.

8. Following denaturation, add 2 µl of 10 × loading buffer (50% glycerol; 1 mM EDTA, pH 8.0; 0.25% bromophenol blue⁷ to the samples and mix thoroughly by micropipetting. Load samples onto the gel immediately.

Note: Because some condensation may form on the inside of the tube during the denaturation incubation period be sure to briefly pulse-centrifuge the tube just prior to the addition of the loading buffer.

9. Electrophorese the samples at a maximum of 5 V/cm distance between the electrodes. Continue electrophoresis until the bromophenol dye front has migrated to within 1 to 2 cm from the distal edge of the gel.

Note: It is especially important to stop the electrophoresis at this point when one is evaluating RNA species of unknown size. In subsequent experiments, RNA gels may be electrophoresed for shorter or longer periods to accommodate specific needs.

10. At the conclusion of the electrophoresis, the gel may be stained in a solution of 0.5 µg/ml ethidium bromide (EtBr; in sterile water or in a fresh aliquot of 1 × MOPS buffer) or, preferably, 1 × SYBR Green II or 1 × SYBR Gold (in 1 × TAE buffer) in order to visualize the RNA within the gel. Staining techniques and options are discussed later in this chapter.

Note: One phenomenon commonly associated with the EtBr staining of formaldehyde gels is unacceptable high background fluorescence, to the extent that the RNA itself may be obscured, due to the presence of formaldehyde. Such gels take on a strong orangepink hue and quite often any nucleic acid in the gel is almost completely obscured. Removal of the formaldehyde requires extensive destaining in two or three changes of nuclease-free H₂O or 5X SSC for 10–15 min each.

11. Regardless of how the gel was stained, it should be soaked in 5 × SSC or nuclease-free H₂O to remove formaldehyde from the gel. Typically, a 20–30 min soaking in several changes of buffer or sterile water is adequate for formaldehyde removal from the gel. Alternatively, the gel may be stored in 300 ml sterile water overnight at 4°.

12. The gel is now ready for photodocumentation.

Urea denaturation

The traditional compounds used to ensure RNA denaturation are toxic, cumbersome, or both. Urea is widely used in polyacrylamide gels for DNA and

⁷The tracking dye xylene cyanol can be included in this loading buffer recipe to a final concentration of 0.25%. This is not recommended, however, because xylene cyanol often interferes with the visualization of RNA (and DNA) after electrophoresis. This is particularly true when a band of interest lies directly under the xylene cyanol.

RNA analysis, and can also be used easily in conjunction with agarose gels. Rather than preparing an agarose gel with urea in it (Locker, 1979), one may treat RNA with freshly prepared urea instead, briefly heating to denature it, and then running the RNA on a non-denaturing agarose gel. This approach is used routinely in this laboratory to assess the integrity of an RNA sample prior to use in RT-PCR, nuclease protection, or other transcription assay. This method also denatures RNA completely enough to support Northern analysis. As shown in Fig. 9.2, a side-by-side comparison of formamide-denatured RNA and urea-denatured RNA shows similar levels of sensitivity, even though the urea-denatured sample shows a greater mobility overall by comparison with the samples that are denatured with the formamide. While it is true that heating RNA briefly without *any* denaturants may be adequate for simply ensuring

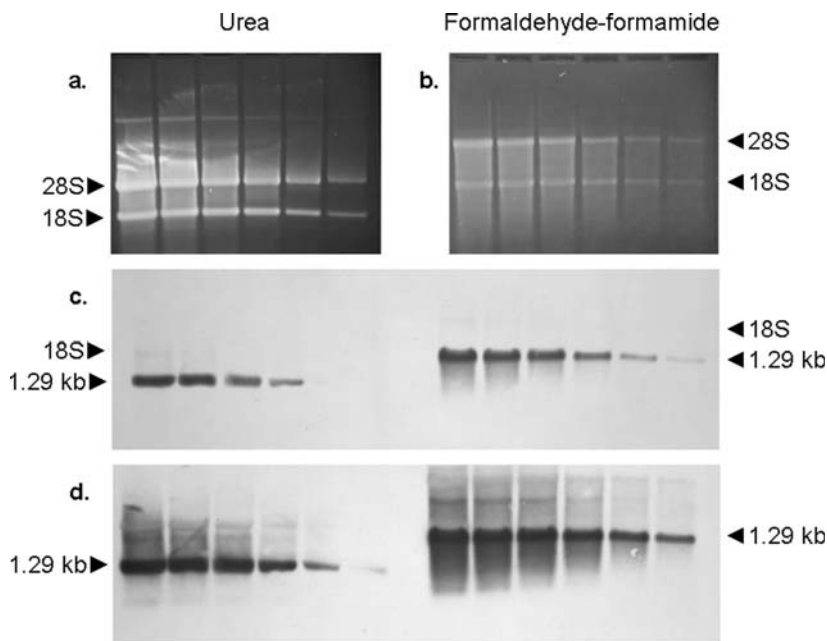


Figure 9.2 Comparative denaturing electrophoresis. Identical amounts of RNA were loaded into each lane: 7.7 μ g (lane 1), 5.8 μ g (lane 2), 3.9 μ g (lane 3), 1.9 μ g (lane 4), 1.0 μ g (lane 5), and 0.5 μ g (lane 6). Panel a: RNA was denatured with urea, electrophoresed in a non-denaturing gel, and then stained with SYBR Green. Panel b: RNA was denatured with formaldehyde and formamide, electrophoresed on an agarose gel containing formaldehyde, and then stained with SYBR Green. The RNA has greater mobility and less background fluorescence (greater contrast) when denatured with urea. Both gels were then (Northern) blotted on the same filter paper and hybridized with a GAPDH probed. Panel c: Detection of GAPDH mRNA via chemiluminescence detection after a 5 min exposure. Panel d: Same filter as in panel c after a 20 min exposure. In both cases, the bands are sharp clear and strong and the GAPDH is in the expected position relative to the ribosomal 28S and 18S species.

sample integrity, the extra minute required to add urea to the sample(s) is easily justified by the greatly improved appearance of RNA on the gel. Urea denaturation is an effective, high-resolution method for rapid assessment of RNA sample integrity without the need for fairly toxic compounds.

Protocol: urea denaturation

1. Prepare a 1.2% agarose gel with $1 \times$ TAE.
2. Prepare a 4M urea stock solution (9.6g in 40ml H₂O). This should be prepared shortly before use.
3. Prepare loading buffer by mixing 100 μ l 10 \times standard gel loading buffer⁸ and 100 μ l fresh prepared 4M urea. The resulting 200 μ l mixture is a 10 \times urea loading buffer stock.
4. For each 10 μ l of RNA in aqueous solution, add 1 μ l of urea loading buffer.
5. Heat to 55 for 10 min.
6. Load gel and electrophorese as usual.
7. At the conclusion of the run, stain the gel in 2 \times SYBR Green in 1 \times TAE buffer and then photograph.
8. Depending on the application the gel may then be used for Northern transfer or simply discarded after confirming sample integrity.

Glyoxal/Dimethyl sulfoxide denaturation

Another common method of denaturing RNA in preparation for electrophoresis is to treat the sample with a mixture of 1M glyoxal and 50% (v/v) DMSO at 50° for 1h (Bantle *et al.*, 1976; McMaster and Carmichael, 1977). Glyoxalation (Salomaa, 1956) introduces an additional ring into guanosine nucleotides and, in so doing, sterically interferes with G:::C base pairing (Hutton and Wetmur, 1973). At high concentrations, glyoxal reacts with all bases of both RNA and DNA (Nakaya *et al.*, 1968); the guanosine-glyoxal adduct is the most stable (Broude and Budowsky, 1971; Shapiro *et al.*, 1970). The preparation of RNA in this fashion precludes some of the potential health hazards associated with the use of formaldehyde. RNA samples are denatured with glyoxal prior to electrophoresis although, unlike the formaldehyde system, glyoxal is not added to the gel. While the glyoxalation process has advantages over formaldehyde denaturation, the advance preparation of glyoxal and required subsequent handling of denatured samples and Northern filters is a major disincentive and the reason why investigators use the formaldehyde denaturing system.

Glyoxal oxidizes very rapidly; therefore, glyoxal stock solutions (40% glyoxal = 6M) must be deionized to neutral pH before use (see Appendix H). If not deionized, the oxidation product, glyoxylic acid, will cause fragmentation of the RNA sample. For the sake of convenience, small aliquots of deionized

⁸Typical 10 \times loading buffer = 50% glycerol, 50mM EDTA (dilute 500mM stock), 0.2% bromophenol blue.

glyoxal can be stored at -20° in tightly capped tubes. Thawed aliquots of deionized glyoxal should be used only once, and the unused portion discarded.

Glyoxalated nucleic acids exhibit decreased electrophoretic mobility compared to an identical sample denatured with formaldehyde. In return, the banding of the glyoxalated RNA tends to be sharper than is observed in formaldehyde-denaturing systems. Further, glyoxalated DNA and RNA molecules are electrophoretically equivalent (McMaster and Carmichael, 1977); therefore, DNA molecular weight standards can be used along side RNA samples if both have been glyoxalated.

The major drawback of the typical glyoxal gel is that the buffering capacity of the standard running buffer, 10mM sodium phosphate, pH 7.0, is quite poor. Phosphate buffer has been used in the past because low ionic strength electrophoresis buffers at pH 7.0 maintain the maximum glyoxal denaturation effect and preclude the need for adding a denaturant to the gel itself. The inability of the sodium phosphate buffer to resist radical changes in pH, however, necessitates constant recirculation of the running buffer throughout the electrophoresis period to maintain pH 7.0. This is imperative because glyoxal will dissociate from RNA (and DNA) when the pH is greater than 8.0 (Thomas, 1980) and because the RNA sample is susceptible to alkaline hydrolysis. Conveniently, most electrophoresis chambers can be purchased with optional buffer ports needed to support buffer recirculation. If recirculation is not possible, the power supply should be disconnected and the buffer changed every 30 min during the run, though this approach is usually quite unsatisfactory. Alternatively, glyoxal gels may be electrophoresed in $1 \times$ MOPS buffer, rather than phosphate buffer, thereby eliminating the need to recirculate the buffer. If running an RNA gel for the first time, the information at the end of this chapter may be very useful.

Protocol: glyoxalation and electrophoresis of RNA

In advance: phosphate buffer preparation

Sodium phosphate electrophoresis buffer is prepared in advance by combining stock solutions of NaH_2PO_4 (monobasic) and Na_2HPO_4 (dibasic) in the appropriate ratio to yield a solution of the desired pH (Table 9.1). The resulting sodium phosphate buffer is often described chemically as $\text{Na}_x\text{H}_x\text{PO}_4$, referencing the inclusion of both the monobasic and dibasic salts in the formulation of this buffer.

Prepare 0.2M stock solutions of Na_2HPO_4 and NaH_2PO_4 , and then autoclave or filter sterilize. Next, combine Na_2HPO_4 and NaH_2PO_4 in the proportions indicated in Table 9.1 to produce the desired pH (7.0) of the final mixture. Then add 90% of the remaining water required to dilute the mixture. Check the pH of the diluted phosphate buffer. It is almost always necessary to make minor pH adjustments: add NaH_2PO_4 to lower the pH or Na_2HPO_4 to raise the pH. Important: fine-tuning the pH must be performed with phosphate buffer that has been diluted to the desired final concentration, rather than using the 0.2M stock solutions. Finally, add sterile H_2O to produce the final volume.

Gel and sample preparation

1. Wear gloves!
2. Melt agarose (final concentration of 1.2%) in an appropriate volume of 10 mM $\text{Na}_x\text{H}_x\text{PO}_4$, pH 7.0.
3. When molten agarose has cooled to 55–60°, cast a 12 cm × 14 cm gel and allow to solidify.

Note: Do not add ethidium bromide to glyoxal gels, as this dye will react with the glyoxal. The inclusion of SYBR Green and SYBR Gold should also be avoided.

4. Denature the RNA in an autoclaved microfuge tube by adding:

- 3.7 μl RNA (up to 10 μg per lane)
- 2.7 μl 6 M glyoxal (freshly deionized)
- 8.0 μl DMSO (molecular biology grade)
- 1.6 μl 100 mM $\text{Na}_x\text{H}_x\text{PO}_4$, pH 7.0.

These proportions will render final concentrations of 1 M glyoxal, 50% DMSO, and 10 mM $\text{Na}_x\text{H}_x\text{PO}_4$.

5. Close the tube tightly and incubate the mixture at 50° for 1 h.
6. After the gel has solidified, immerse it in running buffer (10 mM $\text{Na}_x\text{H}_x\text{PO}_4$, pH 7.0). This should be done while the RNA is being glyoxalated.

Note: Be sure to use the smallest amount of buffer necessary to completely cover the gel, including the edges which tend to “smile” by clinging to the side of the casting tray. Carefully remove the comb by slowly pulling it upward with uniform force. Covering the gel with buffer prior to removal of the comb will lubricate the wells and reduce the vacuum effect created when the comb is pulled. If the gel is not immersed in buffer, comb removal may damage the bottom of the wells.

7. At the conclusion of the incubation period, cool the mixture from step 5 to 20° and then add 2 μl of loading buffer (50% glycerol; 10 mM $\text{Na}_x\text{H}_x\text{PO}_4$, pH 7.0; 0.25% bromophenol blue). If necessary, pulse-centrifuge the tubes for 2–3 s to collect the entire sample at the bottom of the tube.
8. Immediately load each sample onto the gel.
9. Electrophorese at no more than 5 V/cm distance between the electrodes, with constant buffer recirculation. Electrophoresis should continue until the bromophenol blue tracking dye has migrated about 80% along the length of the gel.

Note: It is especially important to stop the electrophoresis at this point when one is evaluating RNA species of unknown size. In subsequent experiments, RNA gels may be electrophoresed for shorter or longer periods, as needed.

10. At the conclusion of the electrophoresis, the gel can be used for Northern transfer without prior staining, though this is not recommended. The gel can be stained in a solution of 0.5 $\mu\text{g}/\text{ml}$ ethidium bromide made up in 100 mM ammonium acetate or in a solution of 2× SYBR Green II made up in 1× TBE buffer.

Note 1: The investigator should be aware of potential interaction between glyoxal and ethidium bromide, which may change its spectral properties (Johnson, 1975; McMaster and Carmichael, 1977), and reduce transfer efficiency.

Note 2: As an alternative to EtBr or the SYBR Green dyes, the gel may be stained in a 30 $\mu\text{g}/\text{ml}$ solution of acridine orange (made up in nuclease-free 10 mM sodium phosphate, pH 7.0). This is not recommended, however, because acridine orange results in notoriously high levels of background in the gel. The required destaining time for acridine orange-stained gels is considerable, and is ideally performed for about

1 h at room temperature in an enameled metal pan (McMaster and Carmichael, 1977) because enamel adsorbs acridine orange. To remove acridine orange from the enamel, rinse with running hot tap water for 10 to 15 min. Polyacrylamide gels are destained in the dark for 2 h at room temperature or at 4° overnight. Handle acridine orange with care because it is a DNA binding dye and, as such, a mutagen.

11. Glyoxalated RNA may be transferred from the agarose gel to a filter membrane, preferably nylon, without any pre-treatment to remove glyoxal. Glyoxal is removed from the RNA during a post-transfer, prehybridization wash of the filter paper (described in Chapter 13).

Note: Soaking the gel in alkali to remove the glyoxal is not recommended, as described in certain older protocols, because such treatment can reduce transfer efficiency and, if not monitored, hydrolyze the RNA to a point at which it is no longer hybridization-competent.

Running RNA on non-denaturing gels

Many attempts have been made to streamline the process of RNA denaturation and, at the same time, reduce or altogether eliminate the use of toxic denaturants. For example, an RNA sample may simply be heat-denatured for a few minutes in a modified loading buffer (e.g. containing SDS) and then immediately electrophoresed on a non-denaturation gel (Kevil *et al.*, 1997). In this regard, some investigators have a preference the TAE buffering system while other investigators prefer running gels buffered with TBE. In some laboratories, investigators report simply heating RNA samples to 65° for 15 minutes in loading buffer without any SDS or other denaturants, and simply running the RNA on the same type of gel that one might ordinary use for the electrophoresis of DNA. While this may be a satisfactory method for some RNA preps, experiences in this lab have shown that not all RNA samples denature as well as others when heated only, and the observed variability seems to be related to the method used to isolate the RNA from the biological source. This is something that will need to be determined empirically in one's own laboratory. If heating alone does not produce a satisfactory electrophoretic profile, the addition of urea to the loading buffer, as described above, usually resolves denaturation issues and the RNA may still be run on a non-denaturing gel. A bit of advice: when working with RNA, keep the agarose concentration between 1.0–1.2% in order to maximize the resolving power of an agarose gel. Finally, keep in mind that, when present, denaturants are usually powerful inhibitors of RNase and may help to protect the sample from degradation.

Molecular weight standards

The accurate determination of the size of RNA species is just as important as deduction of the molecular weight of any other macromolecules subjected to electrophoresis. This is accomplished by comparing the electrophoresis of an

RNA sample against molecular weight standards (also known as size standards or molecular weight markers). Because of the intrinsic chemical differences between DNA and RNA, although both are nucleic acids, the length of RNA molecules increases their rate of migration faster than that of double-stranded DNA size standards of comparable length. This is especially true of agarose gels that contain formaldehyde (Wicks, 1986). Clearly, it is more desirable to electrophorese RNA molecular weight standards in an RNA gel. As noted above, DNA and RNA are not electrophoretically equivalent unless both have been glyoxalated (McMaster and Carmichael, 1977); even then DNAs are useful as size standards only if they fall within the range of the majority of cellular RNAs.

Useful markers for RNA electrophoresis are not at all difficult to come by even in laboratories with minimal funding. There are two basic approaches for the size determination of RNAs of interest:

1. *External molecular weight standards* are nucleic acid molecules of known size that are electrophoresed on the same gel as experimental samples in an unoccupied lane. These standards are commercially available from most molecular biology vendors. For RNA gels, molecular weight standards typically consist of a mixture of 5 to 10 single-stranded RNAs that were produced by *in vitro* transcription. Because these markers are, in fact, single-stranded, they must be denatured by the same method used to denature the samples. Care must be taken to protect RNA molecular weight markers from degradation, which are just as susceptible to RNase degradation as the experimental RNA. Markers that cover various molecular weight ranges are widely available.
2. *Internal molecular weight standards* are RNA species of known size that are actually part of the RNA experimental sample, most often the large and small ribosomal RNAs. As these markers are part of the sample, they will be denatured during preparation of the sample for electrophoresis.

The molecular weight of the ribosomal RNAs are known for many organisms, and their relative positions on a gel are almost always indicated to provide the reader with at least an approximate frame of reference. Whereas more precise measurements of molecular weight are possible using a mixture of RNA size standards, the use of the ribosomal 28S and 18S RNAs⁹ alone or in conjunction with other size standards is nearly universal. Depending on the method of lysis, a properly denatured RNA sample may also show a third, higher molecular weight band upon post-electrophoresis staining. This less obvious transcript is the 45S rRNA, a nuclear precursor to the 28S and 18S rRNAs (Fig. 9.3) that accumulate in the cytoplasm and which form functional ribosomes upon interaction with a plethora of proteins and other molecules.

Proper use of size standards

Either internal or external size standards can be used with confidence on an RNA gel. Relying on one type or the other, or both, has several advantages as

⁹In prokaryotic cells, the 23S and 16S rRNAs are the counterparts of the eukaryotic 28S and 18S rRNAs.

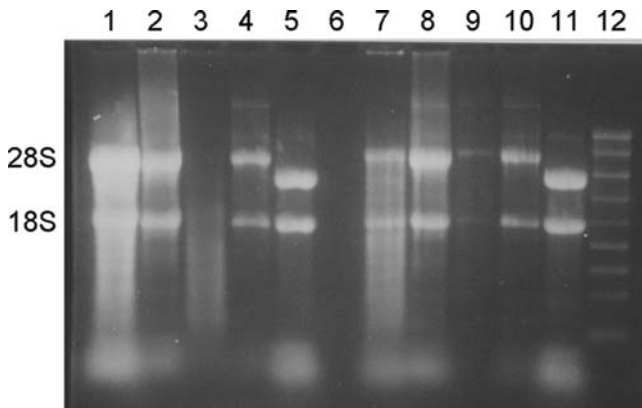


Figure 9.3 Comparison of eukaryotic RNA profiles from different species. Samples of RNA purified from CHO-K1 cells (lanes 1–4 and 7–10) and yeast (*Saccharomyces uvarum*; lanes 5 and 11) on a denaturing 1.2% agarose gel. (Lane 1) Total cytoplasmic RNA, purified by non-ionic lysis – good quality, though the lane is overloaded. (Lane 2) Total cellular RNA purified using guanidinium-acid-phenol – fair quality, with evidence of high molecular weight hnRNA. (Lane 3) Poly(A)⁺ RNA – sample is degraded. (Lane 4) Poly(A)⁻ RNA – good quality. (Lane 5) Total yeast RNA, isolated by acid-phenol – good quality. (Lane 6) Empty. (Lane 7) Total cytoplasmic RNA, purified by non-ionic lysis – poor quality, showing signs of degradation. (Lane 8) Total cellular RNA purified using guanidinium-acid-phenol – good quality, with evidence of high molecular weight hnRNA. (Lane 9) Poly(A)⁺ RNA – good quality. Notice the trace amount of carryover 28S and 18S rRNA that demonstrate the integrity of the sample. (Lane 10) Poly(A)⁻ RNA – good quality. (Lane 11) Total yeast RNA, isolated by acid-phenol – good quality. (Lane 12) Single-stranded RNA molecular weight standard from Promega. Notice the shift (greater mobility) in the 28S rRNA in the yeast sample in lanes 5 and 11, compared to the CHO-K1 samples. Notice also the presence of hnRNA and the precursor 45S rRNA in lanes 2, 4, 8, and 10.

well as drawbacks. With respect to external standards, the key decisions that must be made are:

1. The amount of the standard to use.
2. The positioning of the standard in the gel.
3. The size range of the standard.
4. The method of visualizing the markers (staining vs. end-labeling vs. detection following hybridization).

One characteristic of several brands of pre-made RNA size standards is that the DNA template upon which the transcripts are synthesized is not always removed from the *in vitro* transcription reaction. Further, many of these templates exhibit limited sequence homology with plasmid vectors commonly used for molecular cloning. A possible consequence of this type of construction is manifested when a nucleic acid probe is labeled and used for hybridization without cutting the probe itself out of the vector, or a probe that has

been removed from a vector is still flanked by limited vector sequences that may hybridize to residual template in the marker lane. Because this hybridization is probably an exact match (nucleotide for nucleotide), even subnanogram quantities of template are more than sufficient to generate a strong signal upon hybridization. This has the unhappy consequence of (1) generating very strong detection signals that may bleed over and obscure signals from hybridization events that may have occurred in adjoining lanes; and (2) having previously unseen bands show up in the molecular weight standard lane, causing confusion with respect to size determination of hybridized experimental transcripts. If there is reason to believe that a strong signal might occur, it is judicious to leave a blank lane in the gel between the size standards and the experimental samples, and to use nanogram quantities of the RNA marker. On the other hand, it will be necessary to load about $2\mu\text{g}/\text{lane}$ in order to have sufficient mass to photograph the molecular weight standards after staining the gel.

If no vector sequences are involved or if it is known that the probe is (was) cloned into a vector that does not hybridize to components of the molecular weight standard, then one may confidently load as much as 2 to $3\mu\text{g}/\text{lane}$, a more than sufficient mass for photodocumentation. If vector sequences are involved, as little as 15 to 20 ng of marker will hybridize very efficiently to complementary vector sequences and render sharp bands at the detection stage.

Just as the mass of marker is an important parameter, so is the placement of the marker. Always load the gel asymmetrically. Very little is more frustrating than obtaining a significant banding pattern and not knowing which lane is which. The confusion can be further compounded by the fact that some investigators invert the gel during the transfer, and then fail to maintain the orientation of the filter paper after it is removed from the surface of the gel following the transfer. Were this not bad enough, inversion of the exposed X-ray film during the developing steps adds yet another level of confusion. When the markers are loaded in the center of the gel, the distinction between left and right becomes further obscured. Consistently loading the gel in an asymmetrical fashion will eliminate at least one source of ambiguity. In this lab, the molecular weight standard is always loaded in the last lane, meaning the last well in the upper right-hand corner of the gel (Fig. 9.4).

The range of RNA size standards differs from one source to the next. New systems are best evaluated using a set of markers that spans the range of the RNAs in the sample (500–10,000 bases is a good place to start). After some definition has been given to “typical results” from an experimental system, one may opt to select a preparation of very large or very small standards for greater precision.

Ribosomal RNA

Eukaryotic ribosomes consist of two main components, namely the large and small subunits. The larger of the two is known as the 60S subunit and the smaller is known as the 40S subunit; “S” stands for Svedberg unit, a biochemical



Figure 9.4 Standard nomenclature for referencing wells on a gel. The side of the gel with the wells is furthest away from the investigator who is standing in front of the gel, ready to load the samples. The wells are numbered consecutively, from left to right.

term used to describe the sedimentation rate of a macromolecule. Each subunit consists of a characteristic rRNA and an assortment of proteins. In the course of an RNA extraction, the 80+ ribosomal proteins are stripped away from the rRNAs which, in turn, co-purify with the other cellular RNA species. Removal of the proteins from the 60S and 40S subunits yields the 28S and 18S rRNA, respectively. The much smaller 5S and 5.8S rRNAs, also integral components of eukaryotic ribosomes, are likewise liberated from the intact ribosomes upon removal of protein. In bacteria, the intact ribosomal subunits are known as 50S and 30S, which render the 23S and 16S rRNA, respectively, following protein removal. The highly abundant rRNAs can be very useful as internal controls by:

1. acting as molecular weight markers;
2. giving an indication of the integrity of the sample by visual inspection;
3. providing evidence that an equivalent mass of total RNA has been loaded into each experimental lane. This can be performed by densitometric analysis or through the use of image analysis software. For an historical perspective, see Bonini and Hofmann (1991) and Correa-Rotter *et al.* (1992).

Because rRNA is by far the most prevalent product of transcription (80–85% of cellular RNA), an aliquot of total cellular or total cytoplasmic eukaryotic RNA should electrophoretically resolve into two very distinct, easily observable bands, the 28S and the 18S rRNAs. The appearance of these bands is convincing evidence of the integrity of the sample, that is, that the sample has not been degraded. In high quality mammalian RNA, the intensity of the 28S rRNA is about twice that of the 18S rRNA; it is interesting to note that this is not the case for all eukaryotes. The ratio approaches 1 as one moves down the evolutionary ladder or as a sample begins to degrade.

In addition to the 28S and 18S rRNAs, an intact RNA sample manifests its mRNA component as a significantly lighter smear that appears slightly above, between, and below the rRNAs. This is the normal appearance of cellular mRNA because of its extremely heterogeneous nature and because the mRNA is usually less than 3% of the total mass of RNA in the cell. Often, in order to achieve maximum resolution of larger molecular weight RNAs, electrophoresis is allowed to continue to the extent that the small 5S rRNA, 5.8S rRNA, and tRNA species actually run off the distal edge of the gel and into the running buffer. This is acceptable because the 5S and 5.8S rRNA species are too small to be useful as molecular weight markers for cellular RNA, as they generally migrate through the gel along with the tRNA at the leading edge of the electrophoretic separation, in the 300- to 400-base range. Although this approach is certainly not recommended when evaluating a new system, after a while it will become clear to the investigator whether resolution can be improved by extending the electrophoresis time.

The electrophoretic profile of plant RNA from green tissue is often quite different from that observed upon electrophoresis of animal cells. In addition to the large and small rRNAs a number of rather abundant other bands are visualized upon staining. These are chloroplast transcripts (Fig. 9.5), the presence of which is taken as an indicator of the integrity of the sample. As long as there is little or no smearing in the lower half of the lane, the RNA will probably perform well in any of a number of downstream applications. While the stability of certain plant transcripts is legendary, it is wise to run a gel just to be sure that the RNA is intact, especially if more than six months have elapsed since the RNA was last examined.

Samples of oligo(dT)-purified poly(A)⁺ RNA almost always manifest trace amounts of carryover rRNA, the banding intensity of which may or may not support size calibration. The very fact that the rRNAs are visible at all is strong evidence of the integrity of the sample. One common mistake made during the selection of poly(A)⁺ mRNA is the discarding of the poly(A)⁻ material, which consists primarily of rRNA and tRNA. If this fraction were reprecipitated and electrophoresed along with experimental samples, the poly(A)⁻ fraction would provide suitable size standards (presence of rRNA) and also act as an excellent negative control because the *absence* of poly(A)⁺ mRNA should allow the investigator to demonstrate that hybridization signals observed in other lanes are not artifactual or indicative of non-specific hybridization. This is especially important when the RNA species of interest are located very close to either the 28S or 18S rRNA. An excellent gel-loading strategy is to run, in order, total RNA, poly(A)⁺, and poly(A)⁻ RNA, as is frequently performed in this laboratory (Fig. 9.6).

Samples in which there has been almost any degree of degradation usually fail to manifest the characteristic formation of the 28S and 18S rRNAs and light smearing of the mRNA. Obliterated samples appear as heavy smears which are localized almost entirely below the level at which the 18S rRNA bands in intact samples (Chapter 6, Fig. 6.4b). Occasionally, the investigator will observe fairly uniform, heavy smearing along most of the length of a lane

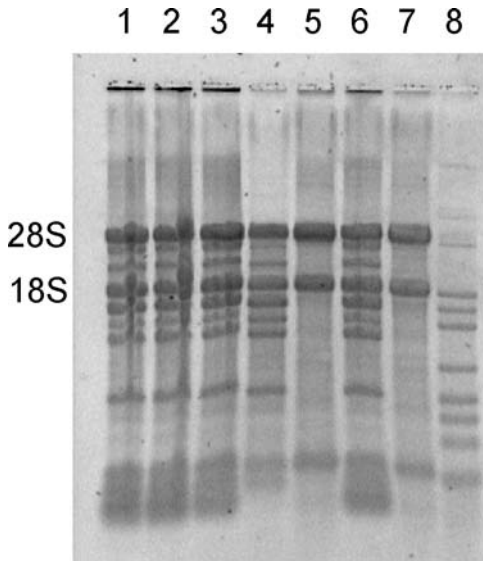


Figure 9.5 Electrophoretic profile of total cellular RNA from Morning Glory, *Ipomoea* sp. Ten micrograms of RNA from each tissue was subjected to standard denaturing electrophoresis at 50V for 3 h. The gel was stained for 30 min in SYBR Gold. Lane 1: First leaf (induced). Lane 2: First leaf (vegetative). Lane 3: First leaf (induced). Lane 4: Cotyledon (induced). Lane 5: Root (induced). Lane 6: First leaf (vegetative). Lane 7: Root (vegetative). Lane 8: RNA molecular weight standard (Promega). In addition to the 28S and 18S rRNA, lanes 1–4, and 6 (RNA from aerial tissue) manifest a number of very abundant chloroplast-specific transcripts, a normal feature of RNA isolated from green tissue in plants. In contrast, the RNA samples from root (lanes 5, 7) exhibit only the 28S and 18S rRNAs. Remarkably, the RNA seen here had been stored at -80° for more than 16 years at the time of electrophoresis (isolated in 1986 by Debra Mohnen) This high quality RNA was subsequently used successfully in a transcript start site mapping study.

on a gel. Within this smear the 28S rRNA and 18S rRNA are usually apparent, seeming to be in the correct positions in the lane, albeit “behind” the smear. There are three possible explanations for an RNA sample with such an appearance: (1) the RNA was not denatured properly; (2) the aliquot is not completely dissolved in buffer; or (3) the RNA is dissolved but the concentration is too high. Failure to completely denature an RNA sample is the most common occurrence which may involve incompletely dissolved RNA. This occurs (1) when large amounts of very pure RNA are precipitated and *not* dissolved in a sufficient quantity of sterile H_2O or buffer to produce a concentration of $5\mu\text{g}/\mu\text{l}$ or less; or (2) when a small-to-moderate amount of RNA was air-dried or dried in a Speed Vac and has not/will not go back into solution¹⁰.

¹⁰Obstinate RNA samples can be “helped” back into solution by heating for 15 min at 65° , by adding more water, or both. In order to maintain the integrity of the sample and its utility in downstream applications, be sure to avoid heating the sample repeatedly. Whatever process is decided upon, it is wise to test it first on a small aliquot of the RNA, rather than risking the entire sample.

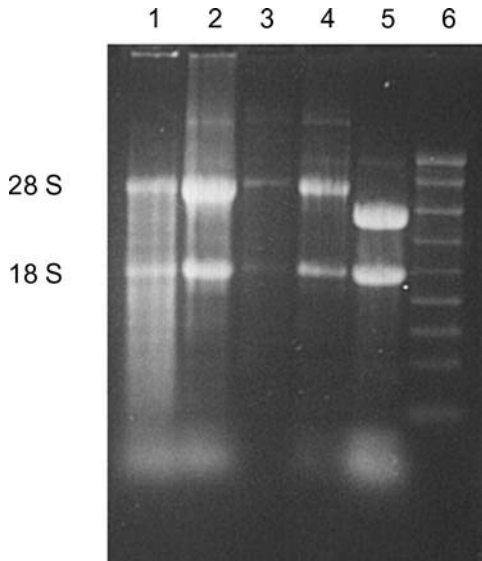


Figure 9.6 Comparison of total cellular RNA, poly(A)⁺ RNA, and poly(A)⁻ RNA. Samples of RNA from CHO-K1 cells (lanes 1–4) and yeast (*Saccharomyces uvarum*; lane 5) were electrophoresed on a denaturing 1.2% agarose gel. (Lane 1) Total cytoplasmic RNA, purified by non-ionic lysis – poor quality. (Lane 2) Total cellular RNA purified using guanidinium-acid-phenol – good quality, with evidence of high molecular weight hnRNA. (Lane 3) Poly(A)⁺ RNA – good quality, with traces of the major rRNA species. (Lane 4) Poly(A)⁻ RNA – good quality. (Lane 5) Total yeast RNA – good quality. (Lane 6) RNA molecular weight standard from Promega. The 28S rRNA in yeast is smaller than in mammals, and thus has greater mobility, compared to the electrophoretic migration of the 28S transcript in the CHO-K1 samples.

Another bit of information that can be conveyed by examining an aliquot of RNA comes from the appearance of the well into which the sample was loaded. Fluorescence coming from within the well suggests that genomic DNA is present in the sample. Enormous fragments of chromosomal DNA, generated by shearing forces often associated with the mechanics of RNA isolation, are unable to enter the gel during the course of the electrophoresis (Fig. 9.7). Further, because UV spectrophotometry does not distinguish between pure samples of RNA and DNA-tainted samples, normalization based on A_{260} may be compromised due to the contribution to the total mass of the sample made by genomic DNA. When other than optimal conditions are suggested by the appearance of a sample on a gel, it is wise to determine the full extent of the problem as well as the quality of the remaining RNA before going on to any subsequent experiments involving a dubious preparation. If DNA tainting is a chronic problem, the symptoms are alleviated by routinely incorporating an incubation with DNase I (RNase-free, of course), as described in Appendix F.

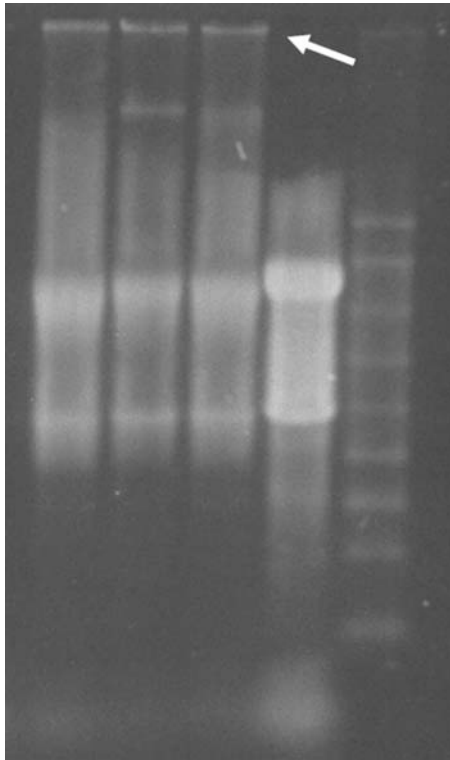


Figure 9.7 Presence of contaminating genomic DNA. The fluorescence apparent in the wells of this RNA gel (white arrow) are indicative of high molecular weight DNA that has carried over from the RNA isolation protocol. DNA can be removed from can RNA sample by incubation with DNase I, as described in Appendix F.

At this point it should be clear that running a gel, even a minigel, is the single best diagnostic available for assessing the integrity and probable utility of a sample.

There is no reason why commercially available size standards cannot be used along with internal rRNA size standards. Having knowledge of the sizes of the 28S and 18S rRNAs, one need only to measure the distance that each of these species has migrated; recall that only two points are needed to draw a straight line and construct a reasonable size calibration curve; these days image analysis software instantly performs these measurements from a Polaroid photograph of a gel or by video capture of a stained wet gel. If image analysis software is not available, it is quite useful to place a fluorescent millimeter ruler somewhere on top of the gel to facilitate measurements, and one should ensure that the placement of the ruler is consistent from one gel to the next. Thus, the UV light is used to generate an image (e.g. SYBR Green or Gold, ethidium bromide), will also cause the fluorescent ruler to appear in the picture as well (Chapter 2, Fig. 2.4). In all cases it is essential to have photographic documentation of even the most preliminary data.

Ultimately, a size calibration curve can be constructed by plotting the \log_{10} of the sizes of the RNA standards against distance migrated from the origin (the well into which the sample was loaded). A curve generated in this fashion can then be used to ascertain the sizes of RNA species observed directly in the gel or that may subsequently be detected by nucleic acid hybridization. Size determinations are now a standard feature of all image analysis software and workstations. Alternatively, a reasonable size calibration curve can be constructed manually (circa 1980) with a couple of data points, a ruler, and a sheet of semi-log paper.

The utility of the ribosomal species as markers is completely dependent on knowledge of the sizes of these RNAs either from the literature or from prior empirical determination. It is unwise to proceed in this manner unless the molecular weight of the ribosomal RNAs is known for the species under investigation. There is considerable variation in the sizes of the rRNAs even among the mammals and, in particular, with respect to the 28S rRNA. There is much less variability observed among 18S rRNAs. In general, 18S rRNAs range from 1.7 to 1.9 kb, while the range of 28S rRNAs is about 4.6 to 5.2 kb¹¹. As a supplementary tool, the location of the tracking dye, most commonly bromophenol blue, may be used as an indicator of the progress of the migration of the samples. While performing electrophoresis, bromophenol blue generally runs with the leading edge of the migrating sample. Another dye, xylene cyanol, runs just behind the 18S ribosomal RNA subunit in 1.2% agarose gels. Xylene cyanol is not used routinely in this lab because of its inconvenient co-migration with the 18S rRNA, and 500 to 600 bp PCR products, often obscuring the fluorescence coming from the bands of interest. Succinctly, the use of xylene cyanol confers no practical advantages and only makes quantification more difficult. Get rid of it.

When characterizing a sample by Northern analysis for the first time it is judicious to continue electrophoresis just until the bromophenol blue has traveled three fourths along the length of the gel. This will produce very reasonable separation of RNA in medium size gels (12 cm × 14 cm), which will simplify accurate size determination of hybridized species while ensuring that small RNAs of possible interest are not lost over the distal edge of the gel into the running buffer.

The reliability of molecular weight data is only enhanced when measurements are derived from both internal and external standards, especially if the RNAs of interest are very large or very small. This is true for two fundamental reasons. First, the logarithmic separation of molecules in a gel is not linear over the entire length of the gel. Very large or very small RNAs that end up outside the linear range of the gel and whose sizes are derived from markers that are within the linear portion of the gel may be sized inaccurately. When external markers are used, it is likely that at least one of the multiple species will electrophorese closer to the RNA of interest; if that RNA is outside the

¹¹In humans the sizes of the 28S and 18S rRNAs are 5025 and 1868 bases, respectively.

linear range of the gel, then some of the bands that collectively make up the marker may also run outside the linear range as well. This translates into fewer errors in size determination. Second, one should never rule out the possibility of uneven transfer of sample material during the Northern transfer process. For example, an air bubble trapped between the gel and the filter membrane or between the gel and the wicking material will drastically reduce the local transfer of sample. In one unfortunate scenario, several samples purportedly exhibit differential modulation of a gene of interest following post-hybridization detection. In reality, signal modulation among the samples may be the result of uneven transfer: potential target sequence molecules may have been left behind in the gel! In order to assess whether efficient and complete transfer has occurred, one or more of the following may be useful:

1. At the conclusion of the blotting period, the gel can be stained again (EtBr; SYBR Green) and inspected for residual RNA, especially rRNA (there should be none).
2. The membrane itself can be reversibly stained with methylene blue to identify directly the positioning of the ribosomal 18S and 28S rRNAs and, in so doing, serve to monitor transfer efficiency.
3. If the gel was stained prior to transfer and small amounts of EtBr or one of the SYBR stains remained in the gel and associated with the sample, then visualization of the rRNAs directly on the filter membrane may be possible if the filter paper is briefly placed on a transilluminator or observed by overhead irradiation with a low-intensity, handheld UV light source.

Remember: the rRNA species serve as excellent gel loading controls as well as transfer controls.

Gel staining techniques

Ethidium bromide

Ethidium bromide¹² (EtBr) remains a widely used dye for the visualization of nucleic acids in agarose gels. It is a phenanthridinium intercalator, structurally similar to propidium iodide. EtBr binds DNA with no apparent sequence preference once every 4–5 base pairs (Waring, 1965). In view of the emergence of SYBR Green and SYBR Gold as realistic alternative methods for staining (discussed below), EtBr is really no longer the best choice. In addition, the issue of *when* to apply the dye is a source of considerable intralaboratory debate. Not only can EtBr be added to the gel, or used to stain the gel at the conclusion of the electrophoresis, EtBr can be added directly to the RNA sample prior to loading the well(s), and otherwise omitted from the gel itself. This latter approach cuts back on the amount of EtBr waste that is generated, and will limit the fluorescence to the lanes in which a sample is loaded. Briefly, prepare a 1 µg/ml stock of EtBr in nuclease-free water. Add 2 µl of EtBr to 18 µl heat-denatured RNA and mix by pipetting up and down, and then load the gel as usual.

¹²Ethidium bromide = 2,7-Diamino-10-ethyl-9-phenyl-phenanthridinium bromide.

Before electing to use EtBr as a staining tool, it is incumbent upon the investigator to weigh four key disadvantages associated with the use of EtBr into this decision:

1. Ethidium bromide is a powerful mutagen (MacGregor and Johnson, 1977; Fukunaga *et al.*, 1984). Extreme care must be exercised when preparing and manipulating ethidium bromide stock solutions and dilutions thereof. Thus, contaminated buffers must be treated as toxic waste and equipment likewise treated as such; EtBr must be inactivated and disposed of properly (Appendix E; Lunn and Sansone, 1987; 1990).
2. The presence of ethidium bromide in a gel or added to samples just prior to loading the gel will retard the electrophoretic migration rate of nucleic acids by about 15%. Although this may not seriously delay the progress of an investigation, it should by no means be used as an excuse to increase the applied voltage to compensate for this phenomenon. Remember that slower-running gels demonstrate enhanced resolution.
3. Continuous exposure to ambient light or, especially, UV light will nick the RNA, making it difficult to work with. Further, EtBr will quickly experience photobleaching, making visualization and quantitative judgments difficult.
4. The presence of ethidium bromide in a gel will reduce its transfer efficiency. If a traditional method of transfer is utilized, such as passive capillary diffusion (Chapter 11), the transfer period may need to be extended for unacceptably long periods. Destaining the gel may facilitate transfer, although even minute amounts of ethidium bromide remaining in the gel can result in poor transfer.

The only advantage of running samples in the presence of ethidium bromide is that the progress of electrophoresis can be monitored by disconnecting the power and briefly irradiating the casting tray upon which the gel is resting by placing it directly on a transilluminator or irradiating it from above with a low intensity UV light source (remember eye and skin protection). If greater separation is required, the gel can be placed back in the gel box and the power supply reconnected for a suitable interval.

If the decision is made to use or continue to use EtBr then the following may be helpful:

1. **Always** wear gloves when handling ethidium bromide contaminated solutions, pipette tips, and so forth. In this laboratory, EtBr-contaminated micropipette tips are placed in conical 15 ml tubes before being placed in the toxic waste container.
2. EtBr is most conveniently prepared as a 10mg/ml stock solution in water and should be stored foil-wrapped at 4°. Microfuge tubes containing stock EtBr should be placed in small storage boxes and also maintained at 4°.
3. Do not exceed a final EtBr concentration of 0.5µg/ml in the gel, the running buffer, or the staining buffer (in instances when the dye is to be used after electrophoresis).
4. Consider staining only the lane(s) containing the size markers and do so after electrophoresis has been completed. To do this, cut the marker lane(s) directly from the gel with a razor blade and stain the isolated lanes as usual. After photographing, the severed molecular weight standard lane and the rest of the gel can be placed next to each other during Northern transfer. The drawback is that one will not be able to assess the integrity of the experimental material unless the membrane is reversibly

stained (described below). If necessary, the nylon filter onto which the RNA is blotted by the Northern transfer process can also be cut, in order to physically distance the lane with the molecular weight standards from other nucleic acid species elsewhere on the gel.

5. If adding EtBr directly to the gel, the molten agarose temperature should first be cooled to 60° because EtBr is heat-labile. In this instance EtBr should also be added to the running buffer; otherwise, the EtBr will be drawn out of the lower portion of the gel as the electrophoresis progresses (Fig. 9.8). Do not forget to dispose of the residual running buffer properly at the end of the electrophoresis period.
6. Do not add EtBr to glyoxal gels, as this dye will react with the glyoxal.
7. Adding EtBr to the samples alone, rather than to molten agarose, will certainly reduce the total mass of the dye required; however, EtBr will still be a transfer obstacle unless it is drawn out of the gel by destaining.
8. Adding EtBr to gels containing formaldehyde can be especially troublesome because formaldehyde gels exhibit very high background fluorescence when stained with EtBr. Such gels often require extensive destaining before even the abundant rRNA species can be observed.
9. Destain gels in the same buffer that was used to stain the gel such as running buffer or nuclease-free H₂O. Alternatively, destain the gel in the buffer that will be used for Northern transfer (5 × SSPE, for example).
10. Make certain that the staining and destaining buffers were at least autoclaved to remove residual RNase. Otherwise, the sample may become partially degraded.

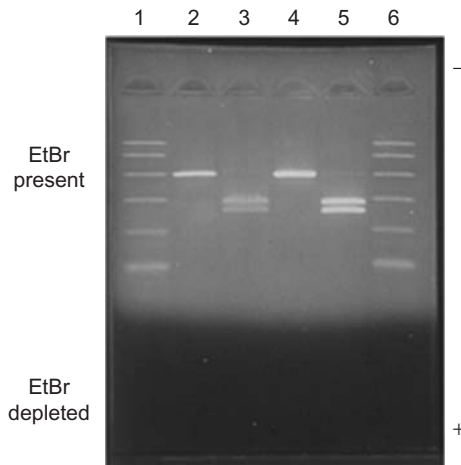


Figure 9.8 Effect of adding EtBr directly to an agarose gel. As the sample runs toward the anode, the EtBr migrates toward the cathode. This has the effect of removing the ethidium bromide from the lower portion of the gel. The longer the electrophoresis continues, the more EtBr will migrate out of the gel. If the dye front happens to pass the sample, some of the fluorescence intensity will be lost, resulting in underestimation of the sample mass. This gel contains a 500 bp PCR product (lanes 2, 4) that has been cut into 273 bp and 227 bp diagnostic fragments (lanes 3, 5). The samples are flanked (lanes 1, 6) by a PCR molecular weight standard (Promega).

SYBR® green

The nucleic acid cyanine binding dyes SYBR Green I and SYBR Green II have become popular stains for in-gel visualization of DNA and RNA, respectively. These dyes are usually provided as a 10,000× stock solution in DMSO. Just prior to use, SYBR Green is diluted to 1× in a Tris-containing buffer, usually 1× TBE, 1× TAE, or TE buffer. Dilution of SYBR Green in H₂O is not recommended because optimal fluorescence occurs in Tris-based buffer around pH 8. Unlike EtBr, the presence of formaldehyde in a gel does not have a negative influence on the background: SYBR Green will fluoresce only when bound to the sample.

Although there are many significant advantages conferred through the use of SYBR Green (Table 9.2), it should not be added to the gel. Samples electrophoresed in the presence of SYBR Green, as opposed to staining the gel afterward, routinely show wave-like, irregular-shaped bands in all lanes and at most molecular weights (Fig. 9.9). This makes mass and molecular weight determinations extremely difficult, if not impossible. In this laboratory, various concentrations of SYBR Green, both above and below the recommended working concentration, have been tested by inclusion in the gel. None have provided satisfactory results. While SYBR Green has been used here extensively with remarkable success, gels are always stained at the conclusion of the electrophoresis. SYBR Green is also compatible with a variety of in-gel enzymatic manipulations and, as needed, can be removed by standard ethanol precipitation.

SYBR® gold

SYBR Gold is currently the most sensitive nucleic acid gel stain. It is a cyanine dye that can be used to visualize single-stranded RNA, double-stranded RNA, single-stranded DNA, or double-stranded DNA in either agarose or polyacrylamide gels, and is comparable in sensitivity to silver staining. In this lab, SYBR Gold has been used successfully in conjunction with a standard 302 nm transilluminator as well as a laser scanner. Like SYBR Green, there is virtually no background fluorescence associated with gel staining, and we routinely stain gels consisting of various agarose concentrations for 30 min under subdued light: we actually cover the gel with an old cardboard box while gently agitating the gel at 50 rpm on an orbital platform. Upon transillumination, DNA and RNA appear as gold fluorescent bands. Like SYBR Green, SYBR Gold is provided as a 10,000× stock solution in DMSO and should be diluted down to a 1× concentration in either 1× TAE or 1× TBE. SYBR Gold is not to be added to the gel prior to electrophoresis for the same reasons for omitting SYBR Green from the gel. Conveniently, the same type of photodocumentation system that has been optimized for SYBR Green-stained gels can be used to photograph gels that have been stained with SYBR Gold.

Table 9.2 Comparison of Ethidium Bromide and SYBR® Stains

	Ethidium Bromide	SYBR Stains
Stock solution	10 mg/ml in H ₂ O	10,000× in DMSO
Working concentration	0.5–1.0 µg/ml in H ₂ O	1× in Tris buffer (pH 7–8) (e.g. 1× TAE, 1× TBE or TE buffer)
Binding affinity for nucleic acids	Moderate to high	Moderate to high
Add directly to gel?	Optional	Should be avoided; always stain gel after electrophoresis
Gel background	High	Low to none
Excitation	300 nm transilluminator	300 nm transilluminator
Color	Pink/Orange	Dignified Green (SYBR Green) Fluorescent Gold (SYBR Gold)
Excitation-induced bleaching of the dye	Yes	No
Filtration	No. 15 deep yellow	No. 15 deep yellow
Limit of detection ¹	5 ng (visual inspection) 600 pg (CCD camera)	60 pg DNA (SYBR Green I) 2 ng RNA (SYBR Green II) 25 pg DNA (SYBR Gold) 1 ng RNA (SYBR Gold)
Mutagenicity ²	High	Less than EtBr; must be handled with the same degree of caution.
Disposal ³	Filtration through activated charcoal	Filtration through activated charcoal
Cost	Low to moderate	Moderate to above-average
Staining time	5–10 min	25–30 min
Destaining helpful?	Yes	Limited benefit

¹Limit of detection for SYBR Green is improved with epi-illumination (Haugland, 1996). Limit of detection can also be improved through the use of a CCD camera or a laser imaging system instead of black and white film.

²Although the SYBR Green I has been shown to be less mutagenic than ethidium bromide (Singer *et al.*, 1999), comprehensive mutagenicity assays have not yet been performed on SYBR Green II or SYBR Gold as of this writing. It is wise to handle *all* DNA binding dyes with care, assuming that each one is mutagenic enough to cause harm.

³The Extractor (Whatman; Cat. No. 10448031) is an excellent, economical device for the decontamination of large volumes of ethidium bromide (and SYBR Green). Treated filtrates are poured down the drain, and the filter itself is discarded according to in-house departmental guidelines. See Appendix E for details.

SYBR® safe

SYBR Safe is a non-toxic alternative to ethidium bromide and the SYBR Green dyes. In fact, SYBR Safe is not classified as hazardous waste under US federal guidelines (Invitrogen web site) and some institutions allow SYBR Safe to be disposed of by pouring it down the drain (check with your safety office first). Of course, common sense dictates that all nucleic acid stains should be handled

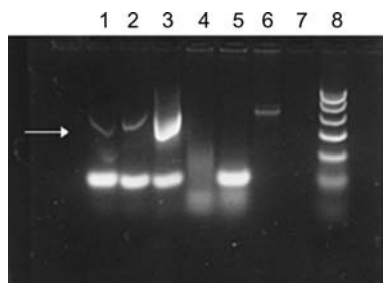


Figure 9.9 Effect of adding SYBR Green to agarose before casting the gel. Both DNA and RNA, at nearly all molecular weights, experience profound band distortion, confounding mass and molecular weight determinations. In contrast, staining gels with SYBR Green or SYBR Gold *after* electrophoresis results in superior staining.

with a high level of caution. Many investigators use SYBR Safe because the annual cost reduction in hazardous material disposal can be substantial.

SYBR Safe is provided as a 10,000 \times concentrate and is diluted in TBE or TAE buffer to a working concentration 1 \times . It is suitable for staining DNA or RNA. Agarose gels may be prepared with SYBR Safe in it or may be stained after the gel has been run. It should be noted that SYBR Safe added to gels behaves like ethidium bromide, meaning that it migrates in the opposite direction of DNA migration. Thus, there may be clearing or loss of fluorescence near the bottom of the gel. Gels stained with SYBR Safe may be viewed with a standard laboratory transilluminator and photodocumented electronically or on Polaroid film.

GelStar[®]

Like the SYBR stains described above, GelStar Stain (Lonza Rockland, Inc.) is a fluorescent nucleic acid stain that offers advantages over traditional staining by ethidium bromide. While the technical data for this product suggests that it can be added directly to the gel prior to electrophoresis or stained afterward, in this laboratory the gels are always stained at the end of the run. GelStar is suitable for staining double- or single-stranded RNA or DNA on either agarose or polyacrylamide gels, and offers an increase in sensitivity compared to identical gels stained with ethidium bromide. GelStar is stable for at least 12 months at -20° when stored desiccated and protected from light. It is provided as a 10,000 \times concentrate in DMSO and functions best when diluted in a Tris-buffered solution between pH 7.0–8.5. GelStar should be handled with the same care as any of the other DNA binding dyes, and disposed of through an activated charcoal filter (Appendix E).

Silver staining

Silver staining is a very sensitive method for detecting small amounts of proteins and low-molecular-weight nucleic acids in polyacrylamide gels. In general, the method is about 100 times more sensitive than Coomassie blue for

many proteins and several times more sensitive than EtBr for visualization of polynucleotides. It has the advantages of allowing visualization of the sample without any specialized equipment and rendering gel backgrounds that are virtually colorless. Silver staining of agarose gels is not recommended mainly because of the drastically reduced sensitivity in this matrix. Briefly, after acid fixation, polyacrylamide gels are impregnated with soluble silver ions (Ag^+), followed by reduction to metallic silver. The initial deposit of the insoluble metallic silver initiates the autocatalytic deposit of more silver, manifested ultimately by visualization of the sample. All of the reagents needed for silver staining are readily available in kits from most major suppliers.

Acridine orange

An older, alternative technique for visualizing nucleic acids molecules in a gel is to stain with the cationic dye acridine orange (McMaster and Carmichael, 1977; Carmichael and McMaster, 1980), historically used in conjunction with glyoxalated RNA and also as a cytochemical stain (Kasten, 1967); these days it is used for cell cycle determination and in the study of apoptosis. Acridine orange does not exhibit qualities that inhibit nucleic acid transfer from gels to the extent that inhibition is observed with ethidium bromide.

Acridine orange will either (1) bind electrostatically to the phosphate groups of single-stranded molecules and fluoresce red at about 650 nm (Bradley and Wolf, 1959); or (2) intercalate into double-helical molecules and fluoresce green at about 525 nm (Lerman, 1961; 1963). Thus, in the same gel, single-stranded nucleic acid molecules appear red/orange, while double-stranded molecules such as DNA and non-denatured RNA appear green. Given the versatility of this dye, the investigator may also detect incomplete denaturation or partial renaturation of an RNA sample before moving on to subsequent experiments. The main reason that acridine orange has not found widespread use is because of the immenseness of the background in gels so stained, and the requirement for unrealistic, heroic efforts to destain the gel.

The staining of agarose gels is accomplished by soaking them in a solution of 30 $\mu\text{g}/\text{ml}$ acridine orange in 10 mM sodium phosphate buffer, pH 7.0 for 30 min. Polyacrylamide gels may also be stained, and in as few as 15 min. To reduce the very high background produced by staining with acridine orange, gels are destained by soaking them in 10 mM sodium phosphate buffer for 1–2 h. It has also been suggested that running hot water over a glyoxal-stained gel for 20 min may help to destain it. It should be easy to detect as little as 0.05 μg of double-stranded nucleic acid and 0.1 μg single-stranded nucleic acid upon UV illumination. Other applications for glyoxal and acridine orange staining have been described elsewhere (Broude and Budowsky, 1971; Carmichael and McMaster, 1980; Sambrook and Russell, 2001).

Methylene blue

Over the past 10 years, it has become fashionable in certain “low-tech” circles to stain agarose gels with methylene blue. While it is true that DNA and RNA

will pick up the stain, the background in the gel is very high, and the sensitivity of the dye is very poor. So, even after extensive destaining, methylene blue has little use as a DNA or RNA stain beyond being able to simply show that a band is present in the gel. As such, methylene blue staining is good for high school labs where toxic compounds such as ethidium bromide and SYBR dyes are not allowed or are too expensive, but methylene blue is not used in the mainstream. A number of companies sell methylene blue-based stains; while it is true that some of these proprietary formulations clearly produce less ambiguous staining than with methylene blue alone, superior sensitivity is achievable with SYBR Green and SYBR Gold.

As another alternative to the approach of staining nucleic acids in the gel with ethidium bromide or with one of the SYBR dyes, Herrin and Schmidt (1988) described a technique by which RNA transferred to nylon or polyvinylidene difluoride (PVDF) filter membranes can be reversibly stained with a solution of methylene blue prior to hybridization.

Advantages:

1. The RNA can be visualized directly on the filter membrane and the locations of the ribosomal RNA species and other size standards identified and marked appropriately.
2. The gel can be stained *after* the transfer has been completed, rather than before, in order to verify that efficient transfer has occurred.
3. The absence of stain in the gel (especially ethidium bromide) prior to blotting will facilitate much more efficient transfer of RNA from the gel.
4. Methylene blue does not pose the potential health hazards and disposal difficulties associated with ethidium bromide and other nucleic acid binding dyes.

Disadvantages:

1. Methylene blue staining does not offer nearly the same level of sensitivity as some of the other dyes, chiefly because the background on the filter tends to be unacceptably high.
2. If the electrophoresed samples are degraded or of sub-optimal purity, this may not be known until after the time and expense associated with setting up the transfer.
3. It is rather difficult to remove completely all of the methylene blue from the filter after it has been stained.

Briefly, RNA that has been transferred to nylon should be UV cross-linked with a calibrated light source to the filter paper prior to staining; alternatively, the RNA can be baked onto the filter paper after staining. RNA may be observed directly on a filter paper within 3 min when placed in a solution of 0.02% methylene blue and 0.3 M sodium acetate, pH 5.5. If RNA was cross-linked to the filter, the stain is easily removed by washing the filter paper in $1\times$ SSPE for 15 min or in a solution of 20% ethanol/80% H₂O for 3–5 min. If the filter was stained before baking the RNA onto it, the stain may be removed from the RNA immediately prior to hybridization by washing the filter in a solution of $0.2\times$ SSPE; 0.1% SDS for 15 min at room temperature, with gentle shaking. Succinctly, methylene blue staining of filters is an interesting, though not very practical, method in the contemporary research lab.

Safety considerations and equipment maintenance

Safety is the prime consideration when working with electrophoresis equipment and a power supply. Never attempt to manipulate the gel, gel box, or power supply while conducting electrophoresis or while the power supply is plugged into an electrical outlet. Nor should any attempt ever be made to modify the design of either a gel box or power supply; any alterations could present a serious risk to both the operator and unsuspecting colleagues in the lab.

Well-designed systems make it impossible to open or remove the lid while maintaining the flow of electrical current. In some designs the lid slides off, breaking electrical contact before the interior of the tank is accessible. This permits the safe inspection of the progress of the electrophoresis. The lid must then be replaced to restore electrical connections and resume electrophoresis. Moreover, it is wise to disconnect the leads from the power supply as well as the gel box when accessing the tank.

Many of the late model power supplies are equipped to accommodate two sets of leads from two different gel boxes. Be sure to leave the power supply off until both of the leads have been connected to the power supply and the gel box. Similarly, be sure to turn the power supply off before connecting a second set of leads (another gel box) to a power supply that is already operational; after all connections have been made properly, the power supply may be turned on again. Be sure to turn the power supply off before attempting to disconnect any of the leads, and do so using one hand only. Last, never take hold of both leads simultaneously; there is a possibility that he who does so may complete the circuit, resulting in electrocution.

Electrophoresis is literally a daily ritual in most molecular biology laboratories. While good quality electrophoresis gel boxes and power supplies can represent an investment of thousands of dollars, it is an unfortunate, if not potentially dangerous, practice to ignore the upkeep of the instrumentation. Basic preventative maintenance of electrophoresis apparatuses is cost effective and ethically compelling. All equipment for performing electrophoresis, including gel boxes and power supplies, come with suggested maintenance procedures that should be followed diligently and the use of this equipment should be guided by common sense.

The proper maintenance and use of electrophoresis gel boxes and power supplies is an important, though often overlooked part of lab safety. It is clearly the responsibility of the operator to inspect wires, connections, and electrodes on a regular schedule *and* immediately prior to use to prevent potentially life-threatening injuries. All power supplies are to be plugged into properly grounded electrical outlets, as per the manufacturer's safety warnings that accompany the purchase of a new power supply, and *never* use an extension cord. Most power supplies are designed for indoor use only, and should never be allowed to get wet. If the instrument does get wet, unplug it immediately and allow it to dry completely, and then contact the manufacturer prior to re-use.

Perhaps the most common problem is the exposure of bare copper wire at the junction of the power cords, electrode leads, and plug ends. This situation poses a severe shock hazard. Replacement cords are usually available at nominal cost from the manufacturer or distributor and it is well worth ordering a few extras when the equipment is initially purchased. Be attentive also for any indication of corrosion and also for cracks in the electrophoresis chamber or the lid itself. To be safe, all power cords, insulation, electrodes, connection nuts, and gaskets should be scheduled for inspection on a weekly basis. For a detailed regimen for electrophoresis apparatus maintenance, see Landers (1990). **All servicing should be entrusted to qualified personnel only.**

Running agarose gels for the first time: a few tips

Essential vocabulary

Casting tray– The portion of the gel box into which the agarose is poured. Gel casting trays are either stationary or removable, depending on the model. The ends of the tray must be sealed, either by taping or through the use of rubber gaskets, until the gel has solidified completely.

Gel box– The apparatus in which the electrodes are mounted and in which the actual electrophoretic separation is carried out.

Loading buffer– A dense glycerol- or sucrose-based reagent that is added to a sample prior to electrophoresis in order to increase the density of the sample. This will allow the investigator to load the wells of the gel after the gel has been completely submerged in running buffer. Loading buffers routinely contain a colored dye, such as bromophenol blue and/or xylene cyanol, the position of which can be used to track the progress of the electrophoresis. Colloquialisms for loading buffer include “sample buffer” and “blue juice”.

Power supply– The electronic device that will generate constant voltage to conduct the electrophoretic separation. The power supply is connected to the gel box after all of the samples have been loaded and the safety cover closed. Only then is the power supply plugged into an electrical outlet, after which it is switched on. Keep in mind that power supplies are potentially very dangerous; investigators using such an instrument for the first time should seek assistance from qualified, experienced personnel.

Running buffer– The electrolyte solution used to conduct the electric current and in which the gel is submerged in order to dissipate heat. A running buffer is useful only to the extent that it resists pH changes and ion depletion during the electrophoresis.

Points to keep in mind

1. There is a potential for RNase contaminating the electrophoresis apparatus, especially the teeth of the comb, while preparing the gel. While nucleases may be inactive in the presence of some denaturants, they can be transferred to and interfere with latter components of an experiment. If nothing else, the comb should be

soaked in a solution of 3% H₂O₂, or in a solution of 0.1% SDS, 50 mM EDTA, and then rinsed with copious amounts of autoclaved H₂O prior to use. Do not use DEPC or 30% H₂O₂, which will severely damage the gel box and may also harm the person using these compounds.

2. A variety of gel boxes are commercially available. Many models are outfitted with rubber gaskets on both sides of the casting tray; this precludes the necessity of taping the tray when pouring the gel (see Appendix L, Fig. 1). In either case, the ends of the tray should be securely sealed so that molten agarose does not leak out.
3. Watch out for agarose boil-over, due to superheating, especially when melting agarose in a microwave. Boiling agarose can cause serious burns. After weighing out agarose, allow it to swell in electrophoresis buffer for 5–10 min, which will reduce the likelihood of boil over. Continue to use caution, however, as boil-over and spillage may still occur.
4. Molten agarose is clear; upon cooling in the casting tray, the gel will take on an opaque appearance. It is important that the casting tray into which molten agarose is poured not be disturbed until it is obvious that the gel has solidified completely.
5. Gels should be cast to a thickness of 0.5–0.75 cm. Gels that are more than 0.75 cm thick will impede transfer efficiency and require exceptionally long blotting periods to elute the entire sample. Staining and destaining thick gels is also rather difficult. To accommodate samples in large volumes, use combs with wider teeth to increase the capacity of the wells. Standard tooth sizes are 1, 1.5, and 2 mm.
6. Make certain that the casting tray is resting on a level surface before pouring the gel.
7. The best way to cast a gel is to pour the molten agarose solution slowly and evenly into the center of the casting tray before putting the comb in place. Stop pouring as soon as the entire surface of the casting tray has been covered. Alternatively, make a permanent mark on the side of the casting tray as an indicator of how much agarose to pour. In so doing, everyone in the lab will make gels the same way. These methods will generate a gel with an appropriate thickness each time. Then insert the comb without delay. Following these steps in this order will yield a very homogeneous gel with no hidden air bubbles. If air bubbles do form, they should immediately be popped or pushed to the side of the tray using a sterile micropipette tip.
8. Agarose solutions containing denaturants should be prepared and cast in a fume hood to minimize exposure and possible health hazards.
9. Be certain to allow enough time for the gel to solidify, and then completely cover it with running buffer *before* attempting to remove the comb. The vacuum forces which are created when the comb is pulled upward are minimized if the gel is submerged in the running buffer. Otherwise, one increases the chances of damaging the bottom of the well. If damage to a well is not detected, the RNA sample will be lost as soon as it is pipetted into the well.
10. If the bottoms of the wells of the gel are repeatedly damaged upon removal of the comb, let the gel “soak” in the running buffer for several minutes to lubricate and loosen the comb.
11. Submerge the gel in only the minimum volume necessary to completely cover the gel. This is because the current will travel the path of least resistance, which is around the gel, rather than through it. This will greatly reduce the electrophoresis rate. However, be sure to look at the edges of the gel that were formed against the sides of the casting tray because agarose has a tendency to “smile” by clinging to the sides of the tray. Failure to completely submerge these portions of the gel can result

in high temperature build-up and melting of the agarose because there is no running buffer covering it to dissipate the heat. Finally, avoid loading samples into the outside lanes; the non-uniform thickness of those lanes will result in non-conforming sample migration (edge effects).

12. When possible, use small sample volumes for electrophoresis. In the context of this chapter, a small volume will cause the RNA to enter the gel in the narrowest cross-section, thereby maximizing resolution.
13. All gel loading buffers contain a dense material such as glycerol or sucrose. Addition of the loading buffer to the denatured RNA sample imparts an added density to the sample and, in so doing, allows it to fall through the running buffer and into well. Therefore, all that is required to successfully load the well is that the investigator merely breaks the surface of the running buffer above the well and carefully extrudes the sample from the micropipette tip (Fig. 9.10). If there is difficulty seeing the well because of reflections from the surface of the buffer, it is useful to place a piece of dark tape or a piece of dark paper under the gel box just beneath the wells.
14. All wells should contain loading buffer of the same composition, i.e., don't mix and match loading buffers because high salt concentrations will cause RNA to move slower through the gel, compared to other samples in the gel. This is also the reason why RNA pellets should be washed extensively with 70% ethanol, as described in Chapter 2. Salt can drastically change the electrophoretic mobility of nucleic acids (Fig. 9.11).
15. Be certain to eliminate any air space from the end of the micropipette tip containing a sample, before attempting to load it into the well of a gel. Injecting air into a well ahead of a sample may result in the complete loss of the sample.
16. If, after loading a sample into a well, the sample floats back out, then alcohol from a previous precipitation or wash step was not completely removed before resuspending the RNA in aqueous. The alcohol will need to be removed by precipitating the sample again followed by 70% ethanol washes and complete drying. As long as the RNA remains in residual alcohol, it will be very difficult to load it on a gel.
17. Do not move the gel box after the samples have been loaded into the wells unless the objective is to create a tidal wave that will displace the samples from the wells.

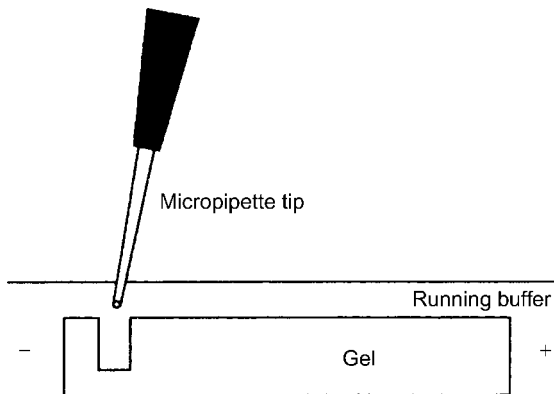


Figure 9.10 Proper positioning of micropipette tip for loading samples into the wells of an agarose gel.

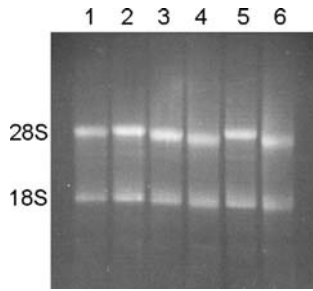


Figure 9.11 Effect of salt on electrophoresis of RNA. Several RNA samples from the same cell culture were isolated simultaneously. At the end of the procedure, the RNA pellets were washed with varying quantities of 70% ethanol. As is expected that the 28S rRNA and 18S rRNA comigrate (line up) from one sample to the next, the obvious differences in electrophoretic mobility are directly related to the fact that the samples in lanes 3, 4, and 6 are in a different ionic-strength environment compared to the samples in lanes 1, 2, and 5. Previously purified RNA can be reprecipitated and washed extensively with 70% ethanol, as needed.

18. When all samples have been loaded into the gel, make sure that the cover of the gel box is secured, to prevent accidental electrocution. Connect the leads to the power supply, and be sure to use gel boxes and power supplies that are appropriately matched. Only then should the power be turned on. Electrophorese at a maximum of 5V per centimeter of gel length. Because the gel can rest in the gel box in either orientation in many models, make certain that the RNA (with its net negative charge) runs from negative to positive (i.e. toward the anode).
19. When inspecting the gel or terminating electrophoresis, be sure to turn the power off before handling either the gel box or the power supply. Further, disconnect power leads one at a time and never handle both at once.
20. Never, ever, attempt to load a gel while the power supply is on, even at minimal voltages. The Author painfully recalls a scientist in a workshop who admitted that “loading the gel at low voltage” improved her data. This is a very dangerous technique that is to be frowned upon and discouraged in every instance.

References

- Alwine, J.C., D.J. Kemp and G.R. Stark. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci.* **74**, 5350–5354.
- Alwine, J.C., D.J. Kemp, B.A. Parker, J. Reiser, J. Renart, G.R. Stark and G.M. Wahl. (1979). Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. *Methods Enzymol.* **68**, 220–242.
- Bailey, J.M. and N. Davidson. (1976). Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Analyt. Biochem.* **70**, 75–85.
- Bantle, J.A., I.H. Maxwell and W.E. Hahn. (1976). Specificity of oligo(dT)-cellulose chromatography in the isolation of polyadenylated RNA. *Analyt. Biochem.* **72**, 413–427.

- Berk, A.J. and P.A. Sharp. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12**, 721–732.
- Boedtker, H. (1971). Conformation independent molecular weight determinations of RNA by gel electrophoresis. *Biochim. Biophys. Acta* **240**, 448–453.
- Bonini, J.A. and C. Hofmann. (1991). A rapid, accurate, nonradioactive method for quantitating RNA on agarose gels. *BioTechniques* **11**(6), 708–710.
- Bradley, D.F. and M.K. Wolf. (1959). Aggregation of dyes bound to polyanions. *Proc. Natl. Acad. Sci.* **45**, 944–952.
- Broude, N.E. and E.I. Budowsky. (1971). The reaction of glyoxal with nucleic acid components III. Kinetics of the reaction with monomers. *Biochim. Biophys. Acta* **254**, 380–388.
- Carmichael, G.G. and G.K. McMaster. (1980). The analysis of nucleic acids in gels using glyoxal and acridine orange. *Meth. Enzymol.* **65**, 380–391.
- Church, G.M. and W. Gilbert. (1984). Genomic Sequencing. *Proc. Natl. Acad. Sci.* **81**, 1991–1995.
- Correa-Rotter, R., C.N. Mariash and M.E. Rosenberg. (1992). Loading and transfer control for Northern hybridization. *BioTechniques* **12**(2), 154–158.
- Farrell, Jr., R.E. and J.J. Greene. (1992). Regulation of *c-myc* and *c-Ha-ras* oncogene expression by cell shape. *J. Cellular Physiol.* **153**, 429–435.
- Favaloro, J.R., R. Triesman and R. Kamen. (1980). Transcriptional maps of polyoma virus-specific RNA: analysis by two dimensional nuclease S1 gel mapping. *Methods Enzymol.* **65**, 718–849.
- Fornace, A.J. and J. Mitchell. (1986). Induction of B2 RNA polymerase III transcription by heat shock: enrichment for heat shock induced sequences in rodent cells by hybridization subtraction. *Nucleic Acids Res.* **14**, 5793–5811.
- Fukunaga, M., B.A. Cox, R.S. von Sprecken and L.W. Yielding. (1984). Production of frameshift mutations in salmonella by phenanthridinium derivatives: enzymatic activation and photoaffinity labeling. *Mutation Res.* **127**, 31–37.
- Gruenwedel, D.W. and N. Davidson. (1966). Complexing and denaturation of DNA by methylmercuric hydroxide. I. Spectrophotometric studies. *J. Mol. Biol.* **21**, 129.
- Haugland, R.P. (1996). "Handbook of Fluorescent Probes and Research Chemicals," 6th edition. Molecular Probes, Inc, Eugene, OR.
- Herrin, D.L. and G.W. Schmidt. (1988). Rapid, reversible staining of Northern blots prior to hybridization. *BioTechniques* **6**(3), 196–200.
- Hollander, M.C. and A.J. Fornace, Jr. (1990). Estimation of relative mRNA content by filter hybridization to a polythymidylate probe. *BioTechniques* **9**, 174–179.
- Hutton, J.R. and J.G. Wetmur. (1973). Effect of chemical modification on the rate of renaturation of deoxyribonucleic acid. Deaminated and glyoxalated deoxyribonucleic acid. *Biochemistry* **12**(3), 558–563.
- Johnson, D. (1975). A new method of DNA denaturation mapping. *Nucleic Acids Res.* **2**, 2049–2054.
- Jungmans, R.P. (1983). A review of the toxicity of methyl mercury compounds with application to occupational exposures associated with laboratory uses. *Environ. Res.* **31**, 1–31.
- Kasten, F.H. (1967). Cytochemical studies with acridine orange and the influence of dye contaminants in the staining of nucleic acids. *Int. Rev. Cytol.* **21**, 141–202.
- Kevil, C.G., L. Walsh, S. Laroux, T. Kalogeris, M.B. Grisham and J.S. Alexander. (1997). An improved, rapid Northern protocol. *Biochem. Biophys. Res. Commun.* **238**, 277–279.

- Landers, T. (1990). Electrophoresis apparatus maintenance. *Focus* 12(2), 54.
- Lehrach, H., D. Diamond, J.M. Wozney and H. Boedtker. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions: a critical reexamination. *Biochemistry* 16, 4743–4751.
- Lerman, L.S. (1961). Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* 3, 18–30.
- Lerman, L.S. (1963). The structure of the DNA–acridine complex. *Proc. Natl. Acad. Sci.* 49, 94–102.
- Lewin, B. (Ed.). (2008). *Genes IX*. Jones and Bartlett Publishers, Boston, MA.
- Locker, J. (1979). Analytical and preparative electrophoresis of RNA in agarose–urea. *Analyt. Biochem.* 98, 358–367.
- Lunn, G. and E.B. Sansone. (1987). Ethidium bromide: destruction and decontamination of solutions. *Analyt. Biochem.* 162, 453–458.
- Lunn, G. and E.B. Sansone. (1990). Degradation of ethidium bromide in alcohols. *BioTechniques* 8(4), 372–373.
- MacGregor, J.R. and I.J. Johnson. (1977). *In vitro* metabolic activation of ethidium bromide and other phenanthridinium compounds: mutagenic activity in *Salmonella typhimurium*. *Mutation Res.* 48, 103–108.
- Maniatis, T., A. Jeffrey and D.G. Kleid. (1975). Nucleotide sequence of the rightward operator of the phage lambda. *Proc. Natl. Acad. Sci.* 72, 1184–1188.
- McMaster, G.K. and G.C. Carmichael. (1977). Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci.* 74, 4835–4838.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn and M.R. Green. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 7, 1175–1193.
- Nakaya, K., O. Takenaka, H. Horinishi and K. Shibata. (1968). Reactions of glyoxal with nucleic acids, nucleotides and their component bases. *Biochim. Biophys. Acta* 161, 23–31.
- Rave, N., R. Crkvenjakov and H. Boedtker. (1979). Identification of procollagen mRNAs transferred to diazobenzoyloxymethyl-paper from formaldehyde gels. *Nucleic Acids Res.* 6, 3559–3567.
- Reijnders, L., P. Sloof, J. Siral and P. Borst. (1973). Gel electrophoresis of RNA under denaturing conditions. *Biochim. Biophys. Acta* 324, 320–333.
- Salomaa, P. (1956). Two volumetric methods for the determination of glyoxal. *Acta Chem. Scand.* 10, 306–310.
- Sambrook, J. and D.W. Russell (Eds). (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, New York.
- Shapiro, R., B.I. Cohen and D.C. Clagett. (1970). Specific acylation of the guanine residues of ribonucleic acid. *J. Biol. Chem.* 245, 2633–2639.
- Sharp, P.A., A.J. Berk and S.M. Berget. (1980). Transcriptional maps of adenovirus. *Meth. Enzymol.* 65, 750–768.
- Singer, V.L., T.E. Lawlor and S. Yue. (1999). Comparison of SYBR Green I nucleic acid gel stain mutagenicity and ethidium bromide mutagenicity in the *Salmonella*/mammalian microsome reverse mutation assay (Ames test). *Mutation Res.* 439, 37–47.
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503–518.

- Spohr, G., M.E. Mirault, T. Imaizumi and K. Scherrer. (1976). Molecular-weight determination of animal-cell RNA by electrophoresis in formamide under fully denaturing conditions on exponential polyacrylamide gels. *Eur. J. Biochem.* **62**, 313–322.
- Thomas, P.S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci.* **77**, 5201–5205.
- Waring, M.J. (1965). Complex formation between ethidium bromide and nucleic acids. *J. Mol. Biol.* **13**, 269–282.
- Wicks, R.J. (1986). RNA molecular weight determination by agarose gel electrophoresis using formaldehyde as denaturant: comparison of RNA and DNA molecular weight markers. *Int. J. Biochem.* **18**, 277–278.
- Zinn, K., D. DeMaio and T. Maniatis. (1983). Identification of two distinct regulatory regions adjacent to the human β -interferon gene. *Cell* **34**, 865–879.

This page intentionally left blank

10 Photodocumentation and Image Analysis

Rationale	221	Traditional methods of photodocumentation	233
Safety first	222	Sample visualization	234
Digital image analysis	223	Filtration	234
Image formats	228	Tips for optimizing electrophoretograms	235
Practical considerations	229	Inherent limitations of photographic and X-ray films	238
Digital image analysis for every budget	231	References and suggested reading	239
Image analysis workshops	232		
PhosphorImagers	232		

Rationale

The absolute necessity for maintaining accurate records of experiments performed and the resulting data is indisputable. Electrophoresis gels contain information critical for accurate interpretation of the outcome of the present experiment; these data often suggest subsequent experiments that may give further definition to a scientific investigation. In both analytical and preparative settings, the information obtained by electrophoresis may be invaluable. Gels deteriorate in a relatively short time, however, and the dyes used to stain nucleic acids in gels deteriorate even more rapidly. This mandates archiving the image electronically and/or producing a tangible photographic record (an electrophoretogram), as well as a photographic record of any associated X-ray films generated upon hybridization detection (by chemiluminescence or autoradiography). One must ensure a stable, long-term record of results for future analysis, for in-house presentations, and for publication.

A variety of systems have gained widespread acceptance and are currently in use to record and analyze the various types of images commonly encountered in a molecular biology setting. These systems fall into two broad categories: (1) digital imaging, in which the image of the gel or X-ray film is analyzed by and stored in a computer and/or on a disk; and (2) the traditional photograph

generated on thermal paper or Polaroid film. Both approaches have numerous advantages and disadvantages. As of this writing, most molecular biology laboratories are equipped with at least one digital image analysis work station. While sophisticated image analysis systems can cost tens of thousands of dollars, electronic archiving of images is possible even in the most modest of laboratories by simply mounting a digital camera above a gel, and by providing the proper filtration for the image and shielding for the user.

Safety first

UV transilluminators and handheld UV light sources are both sources of intense, dangerous ultraviolet radiation. Exposure to UV light from these standard laboratory instruments, which commonly emit UV radiation at 254nm, 302nm, or 312nm, can cause serious damage to the cornea, retina, and other structures of the eye, not to mention the damage to exposed skin. Fortunately, most of the image analysis stations sold today are designed such that the transilluminator is contained entirely within the station, and the UV light will not function unless the door is closed and sealed. This intelligent design prevents the accidental exposure of the user or by-standers to UV light, unless someone foolishly bypasses the safety mechanisms.

For older style photodocumentation systems, one is generally at much greater risk for exposure to UV light because of the basic design of the instrumentation: one must raise the cover of the transilluminator in order to access and photograph the gel, or even to simply point to a specific band for the benefit of a colleague. Because the damage is cumulative, minute amounts of exposure can be catastrophic in both the short- and the long-term. Be sure to wear gloves and protective eye wear/face shield at all times, even when using transilluminators with a UV protective cover, such as the type that covers the gel while the UV light is on. Although these covers usually block most of the UV light, there may still be a small amount of transmission. Moreover, cracks, improper placement, and open areas permitting direct exposure to UV light are all serious potential health and safety risks.

When photographing gels resting on the surface of a UV transilluminator, the gel is usually covered entirely by the hood of handheld cameras. The UV light should be switched on immediately prior to physically taking the picture and should be turned off immediately thereafter, even before removing the exposed film from the camera; because it is so easy to forget that the UV light remains on, you and other personnel in the room could be exposed to dangerous UV light. This occurs frequently when the room lights are on while pictures of the gel are being taken. Exposed sections of the transilluminator surface, no matter how small, pose a great danger to the operator and should be covered at all times; in addition, the transilluminator surface becomes hot very quickly. One should also be aware that not all safety glasses are UV rated; it is critically important to have access to the correct safety equipment when using a UV

transilluminator for any purpose and for any amount of time. Finally, it is important to avoid direct contact with the dyes with which the gels are impregnated, as DNA and RNA binding dyes carry significant mutagenic potential. The dangers associated with photodocumentation are not to be misjudged or underestimated.

Digital image analysis

Many of the traditional methods for the interpretation of photographs of gels (electrophoretograms), including visual inspection and densitometry, have been abandoned in favor of the plethora of computer software and hardware for rapid, reproducible image analysis. Useful measurements that are now automated include determination of molecular weight, concentration, relative abundance, integrated optical density (IOD), and image saturation. The computer also provides a means of cataloging stored images. An image analysis system most often consists of a completely self-contained system, such as the FOTODYNE Luminary/FX workstation (Fig. 10.1) which is in constant use in this laboratory. Alternatively, image analysis software can be purchased separately and used in conjunction with currently existing instrumentation, of which Gel-Pro[®] Analyzer (Media Cybernetics) is a preeminent example (Fig. 10.2). This software uses the point and click approach to derive the desired information,

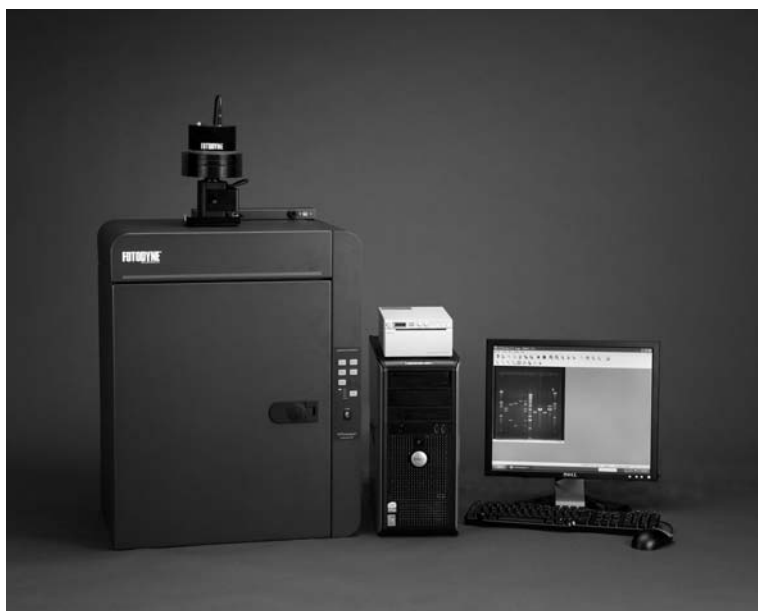


Figure 10.1 FOTO/Analyst[®] Luminary/FX workstation. Courtesy of FOTODYNE (www.fotodyne.com).

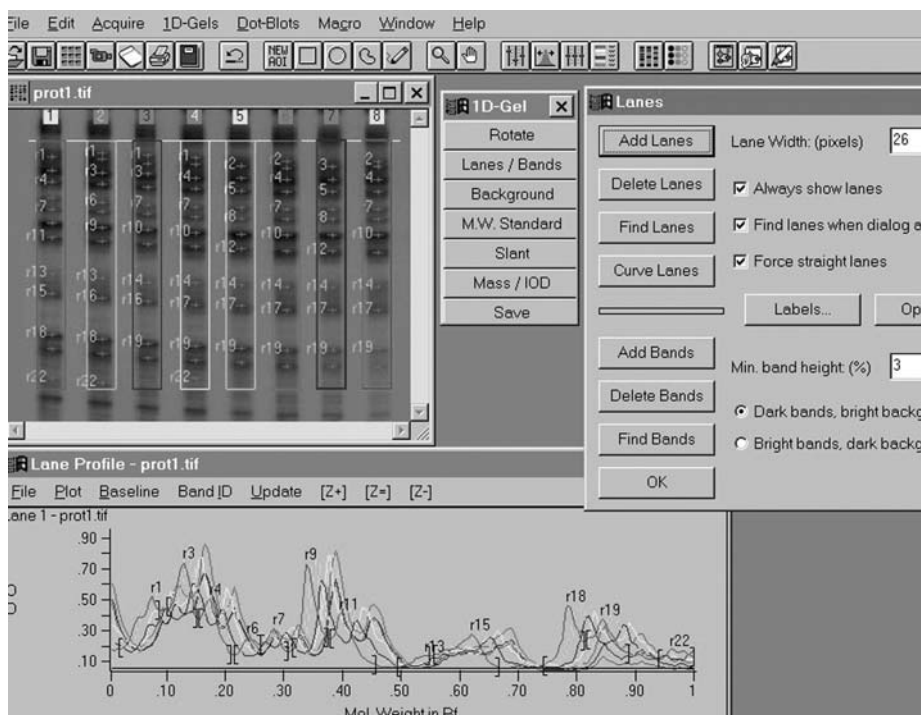


Figure 10.2 Image analysis being performed using Gel-Pro Analyzer. Courtesy of Media Cybernetics (www.mediacy.com).

without demanding computer literacy of the user, and many of the functions, including identifying lanes, finding bands, and performing user-defined macros are automated. Although limited input is required for use, the capacity for complete override of all automated functions allows maximum flexibility.

Meaningful image analysis requires an understanding of the basic terminology and the benefits and deficiencies associated with this approach for analyzing gels and archiving data; thus, a short overview of this technology and practical concerns are presented here.

An image is a visual representation of an object, and image processing manipulates information within an image to make it more useful. Digital image processing is a specific type of image processing performed with a computer. Succinctly, image analysis may be defined as the presentation and quantification of images.

The digitization process divides an image into a horizontal grid, or array, of very small regions, each of which is called a “picture element” or pixel (Fig. 10.3). In the computer, the image is represented by this digital grid, or bitmap. A pixel usually represents a very small region within an image, often 1/300th of a square inch, or less. Each pixel in the bitmap is identified by its position in the grid, referred to by its row (x coordinate) and column (y coordinate) number. By convention,

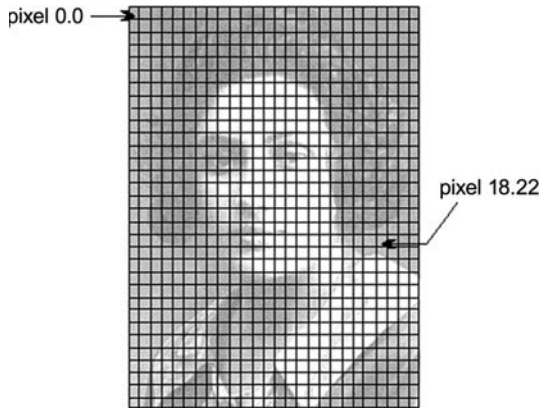


Figure 10.3 Pixel Bitmap. Courtesy of Media Cybernetics.

pixels are referenced from the upper-left position of the bitmap, which is considered position 0, 0 (row 0, column 0).

When an image, such as a photograph of a gel, is digitized, it is examined in grid fashion. This means that each pixel in the image is individually sampled, and its brightness is measured and quantified. This measurement results in a value for the pixel, usually an integer, which represents the brightness or darkness of the image at that point. This value is stored in the corresponding pixel of the computer's image bitmap. When the image is digitized, the width and height of the array is chosen and fixed. Together, the bitmap's pixel width and height are known as its spatial resolution.

Depending on the capability of the measuring hardware and the complexity of the image, anywhere from 1 to 32 bits might be used to store each pixel value. Pixel values for simple images, such as line drawings, containing only black and white information, are represented by a single bit: 0 = black, 1 = white. In contrast, photographic-like images contain much more information; 24 bits are needed to represent all the possible colors that might occur in a true color image. The number of bits used to present the pixel values in an image is referred to as its pixel depth, or bits-per-pixel (BPP). The number of bits-per-pixel used to represent each pixel value determines the image's class.

Although the bit depth (BPP) indicates how many unique colors an image can possess, it does not indicate which colors are actually contained within the image. Color interpretation is determined by bit depth or image class, which include the following:

- Gray scale 8
- Gray scale 12, and
- Gray scale 16.

Gray-scale pixel values represent a level of grayness or brightness, ranging from completely black to completely white. This class is sometimes referred to as "monochrome." In an 8-bit gray scale image, a pixel with a value of 0 is

completely black, and a pixel with a value of 255 is completely white (Fig. 10.4). A value of 127 represents a gray color exactly halfway between black and white (medium-gray), and a pixel value of 64 has a gray color halfway between medium-gray and black. Although gray scale images with bit depths of 2, 4, 6, 12, 16, and 32 exist, 8 BPP gray scale images are the most common because (1) its 1-byte-per pixel size makes it easy to manipulate with a computer; and (2) it can faithfully represent any gray scale image because it provides 256 distinct levels of gray¹.

Often it becomes necessary to enhance an image of a gel in order to maximize the amount of useful information that can be derived from it. Although there are numerous methods for enhancing images, the three most commonly used in digital image analysis are brightness, contrast, and gamma correction.

Brightness is a term used to describe the overall amount of light in an image. When brightness is increased, the value of every pixel in the image is increased, moving each pixel closer to 255, or white. When brightness is decreased, the value of each pixel is likewise decreased, moving it closer to 0, or black.

Contrast is a term used to denote the degree of difference between the brightest and darkest components in an image. An image has poor contrast if it contains only harsh black and white transitions, or contains pixel values within a narrow range. For example, an image with values ranging from 100 to 140 would have poor contrast. An image has good contrast if it is composed of a wide range of brightness values from black to white. The amount of the intensity scale used by an image is called its “dynamic range”. An image with good contrast will have a broad dynamic range.

Gamma correction is a specialized form of contrast enhancement that is designed to enhance contrast in the very dark or very light areas of an image. It does this by changing midtone values, particularly those at the low end, without affecting the highlight (255) and shadow (0) points. Gamma correction can be used to improve the appearance of an image, or to compensate for differences in the way different input and output devices respond to an image. Gamma correction involves applying standard, non-linear γ curves to the intensity scale (Fig. 10.5). A γ value of 1 is equivalent to the identity curve, which has no effect on the image. An increase in the γ value ($\gamma > 1$) will generally lighten an image and increase the contrast in its dark areas. A decrease in the γ value ($\gamma < 1$) will generally darken the image and emphasize contrast in the lighter areas.

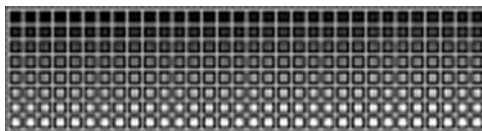
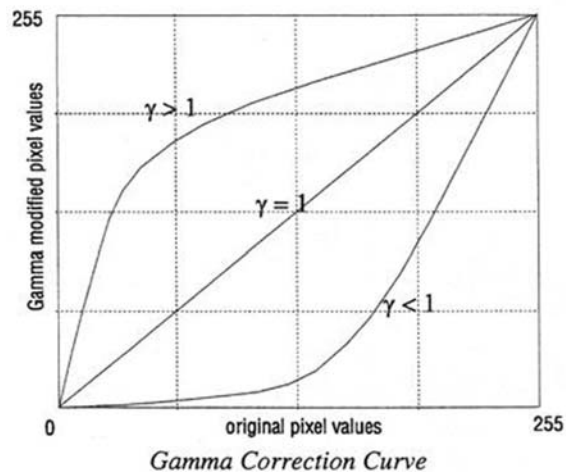
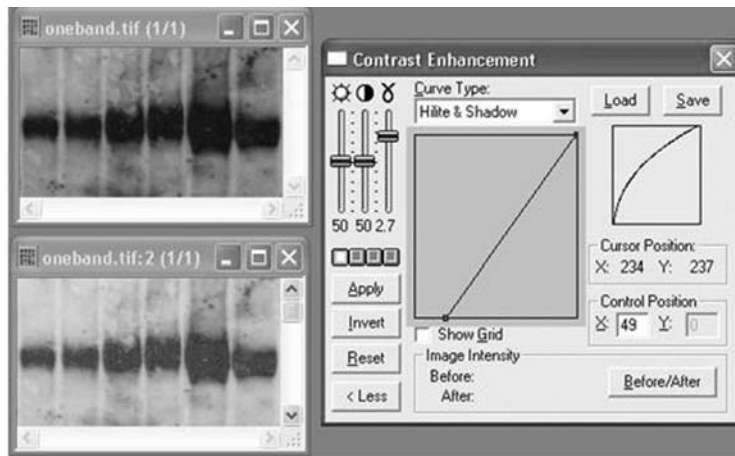


Figure 10.4 256 shades of gray.

¹The human eye can distinguish fewer than 200 gray levels.



a



b

Figure 10.5 Gamma Correction Curve. (a) Effect of γ curves on pixel values 0 through 255. (b) Original Northern blot image (upper panel) and the same image γ correction (lower panel). Note how the applied γ correction of 2.7 has lightened the image while increasing the contrast of the bands. The images were generated using Gel-Pro Analyzer (Media Cybernetics, Bethesda, MD).

Another broad category of operations frequently used to collect data from gel images is known as density analysis. Optical density (OD) analysis is a common image processing application used to determine the amount of matter in a material by measuring the amount of light it transmits (passes through it). Because OD analysis measures the amount of light passing through a material, OD measurements are meaningful only in the analysis of images that have been captured with the light source radiating from *behind* them; for example, gels UV-irradiated from beneath on a transilluminator. Optical density measurements are not useful in the analysis of images captured under reflected light. Molecular biologists are perhaps more familiar with OD in the context of UV spectrophotometric determination of nucleic acid concentration, as described in Chapter 6.

In general, OD of a band or other element of a digitized image assumes an exponential decay of light inside the transmitting material and is calculated as

$$\text{Optical density } (x,y) = \frac{-\log (\text{intensity } (x,y) - \text{black})}{\text{incident} - \text{black}}$$

where:

intensity (x,y) is the intensity at pixel (x,y) ;

black is the intensity generated when no light goes through the material;

incident is the intensity of the incident light.

A material through which no light is transmitted has an infinite optical density, and one through which all light is transmitted has zero optical density. Thus the OD scale is inversely related to the intensity scale, that is, on an optical density scale, dark pixels produce high values, and light pixels produce low values.

Integrated optical density (IOD) is the sum of pixel values minus a background value for each pixel within a user-defined area of interest. In a given area, the increased size of a band or the increased brightness of a spot relates to higher integrated optical density measurements. IODs are very useful when comparing band intensities on images, within or among lanes, in which mass standards have not been included.

Image formats

Digitized images are usually stored on a computer hard drive so that they can always be recalled for further analysis. It is strongly recommended that (1) each user has a personal folder, primarily for organizational purposes; and (2) the entire hard drive should be backed up regularly. Digital images may occupy a large amount of disk space, so it is usually best to crop the image as much as possible, though obviously not to the detriment of the interpretation of the data. When images are stored to a disk, they are stored in a particular format. These formats vary mostly in the types of image data that they support and the types of compression that they make available to reduce the storage size of the image.

Digitized images of gels are most commonly stored as “TIFF” files (*.tif), an acronym for “tagged image file format” (also, “tagged information field format”). Some of the other common image file formats include BMP (bit-map), JPEG (joint photographic experts group), GIF (graphics interchange format), and several others. Once stored, these images can be retrieved as any other file would be, posted on the internet, and transmitted electronically via email.

Practical considerations

The primary concern of most scientists considering the acquisition of image analysis software and hardware is “How much?” Not surprisingly, there is enormous variation in the format of equipment, the variety of functions that can be performed, the system requirements, the amount of supporting hardware required for the system to run and, of course, the cost. As suggested above, many systems are completely self-contained, are elegant to behold, and have a great many functions preprogrammed. Of these, some are expensive and often beyond the budget of many laboratories: at best these might be purchased at the department level and used by everyone. Other digital image analysis products are designed to be loaded directly onto a computer already in the laboratory. Beyond the purchase price of the image analysis software, the only investment required is the purchase of some type of image acquisition device, such as a scanner, a digital (CCD) camera, or an analog camera with a frame grabber. One point that is abundantly clear is that, at all levels of sophistication, the added benefits of image analysis software easily justify the cost (Table 10.1).

A key advantage of using image analysis software is the automation of tedious and time-consuming determinations of the location and mass distributions of molecules in a gel. Calculations commonly associated with the antiquated

Table 10.1 Benefits of Image Analysis Software

Perform standard and customized measurements.
Automate and unify image analysis in the lab.
Save time and increase productivity.
Provide reproducible data analysis ¹ .
Perform intensity and background calibrations.
Compensate for non-uniform background.
Calculate integrated optical density of each band, relative abundance, or mass distribution among bands in the same lane, instead of the error-prone measurements associated with older-style instrumentation.
Export digital image data to Excel or other commonly used spreadsheet software.
Distribute digital images easily by email as TIFF files.
Import digital images easily into Word documents and PowerPoint presentations.

¹This aspect of image analysis is especially important because, upon visual inspection, most people tend to overestimate the integrated optical density of bands on gels and autoradiograms.

approach involve manually generating a size calibration curve and concomitant densitometry for mass determinations. Digital image analysis reduces intra-laboratory error and obviously favors maximum throughput in laboratories generating large numbers of images. Further, better-quality image analysis software is fully compatible with images of gels as well as images on X-ray films and chromogenic detection methods, and is fully able to analyze bands as well as signals in dot blot format. Beyond the classical applications involving analysis of gels, newer versions of image analysis software are in widespread use in conjunction with microarray analysis (Chapter 21) and numerous other cutting-edge applications.

One area of gross misunderstanding about image analysis software pertains to the perceived ability of image analysis software to improve upon the image itself. One must realize first and foremost that if the investigator scans a poor image of a gel, then the image analysis software has a poor image with which to work. If a gel was not electrophoresed for a sufficiently long period, the software cannot further resolve the bands. A good deal of successful image analysis depends on the quality of the image provided by the investigator.

Background correction is one area of acute controversy, particularly when high-precision quantitation of band mass and total lane mass is required. This important aspect of image analysis is not without intrinsic difficulties, the most challenging of which is the fact that the background is never uniform in a gel or on an image. Instead, a gradient of background exists from left to right and top to bottom. As such, there are three basic approaches:

1. Examine numerous small areas in the image, to deal with “local background” only (a so-called “area of interest”).
2. Position the cursor at some neutral location on the gel, sample the background at that point, and then perform a global background subtraction.
3. Position numerous background sampling symbols at various points in the image, and subtract the measured background from the nearest lane in a gel. This method offers the greatest degree of accuracy and versatility for the user.

Finally, photographic films have several inherent limitations, described below. Thus, the potential exists for obscuring the true quantitative differences between bands under investigation on the same gel. Recall that 255 is the largest possible pixel value, and areas within an image containing these values are said to be saturated. While pixel values below 255 will continue to increase with longer exposures, these will also become saturated eventually. Because longer exposure times will not/cannot increase pixel values beyond 255, calculations on bands containing saturated pixels are likely to generate data that are lower than expected, compared to non-pixel-saturated bands. Quantitative measurements are therefore compromised when saturated pixels are involved. The most recent generation of digital image analysis products alert the user, usually by turning a portion of the image red or blue when some or all of the image falls into the saturation range. This allows the user to adjust the image intensity so as to avoid what this Author likes to describe as “saturation-induced misinterpretation”.

Most image analysis systems also allow image integration, a feature that will enhance low intensity bands. Integration of an image occurs by simply capturing the image and then overlaying the same image for as many times as the investigator wishes. To better understand integration, imagine taking a photograph in a low light setting: as a result, many of the details of the subject matter are difficult to see. Now, imagine making an infinite number of transparent copies (referred to as frames) of this same photograph and then overlaying the copies onto the original image one-by-one. This is integration. As a result, the obscured details of an image become apparent, and the investigator may stop the integration process when satisfied with the enhanced appearance of the image. While this is a wonderful feature to have available for image enhancement, it is also very important to understand that the background also increases with the number of applied frames, as do those components of the image that were easy to see in the original image. Integration is a useful tool, but must be used judiciously.

One common mistake that this Author has observed repeatedly is that two different images are with different degrees of integration are compared directly in order to determine the relative abundance of multiple PCR products distributed among the two gels. If one of the gel images was produced with 20 frames of integration and the other gel was produced with five frames of integration, then a direct gel to gel comparison is completely inappropriate. Bands that are to be compared directly for quantification or normalization are to be electrophoresed on the same gel. Alternatively, though less desirable, would be normalization of the bands on one of the gels to a mass standard on the same gel, followed by the same type of correction on the other gel(s) to allow for some degree of gel-to-gel comparison.

Digital image analysis for every budget

Although these systems require a significant up-front investment, it is quite correctly argued that these digital systems offer superior speed and efficiency in the manipulation of data, including the ability to rapidly email images, either as *.tif file attachments or imported into a Word document attachment, to collaborators in distant locations. Polaroid photographs can also be scanned, digitized, recalled, analyzed, distributed, and archived as needed, and traditional photographs have greater longevity than printouts on thermal paper. The point is that a hard copy of each gel image, whether produced on thermal paper or on Polaroid film, is a very important document for inclusion in the laboratory notebook, and there will always be a place for this approach. Computers are convenient, but they also crash from time to time, making a hard copy of an electrophoretogram an absolute necessity. Moreover, from a time-saving perspective, it is much easier to open a lab notebook and look at a print of the gel than to:

- wait for the computer to become available;
- log onto the system;
- open a gel analysis program;
- find the directory in which the image was stored; and
- print out a copy of the stored image, if needed.

The take-home lesson: image analysis software is a highly efficient and time-saving tool, but having a photographic hard copy of a gel is just as important. In this laboratory, digital images are printed out on the very day they are generated so that a hard copy can be placed in a laboratory notebook, in addition to archiving the image on a disk or on a hard drive.

Image analysis workshops

The digital technology revolution has left a number of more seasoned investigators reeling from what might be described as “technology overload” in the area of image analysis approaches and the concomitant interpretation of data. One of the commonly expressed concerns pertaining to the possible purchase of a new digital system in one’s lab is that only one or a few technically savvy individuals will be able to use it! Happily, much of the software in use today does not require an extensive understanding of computers or the mathematics behind the software to get it to work to its full capabilities. Moreover, there are several organizations that provide hands-on image analysis workshops and, in some cases, these organizations will come to you to provide on-site training. The courses are usually one or two-day programs. Some of these workshops are provided at the various society meetings that are held each year, while other tutorials are available on-line. A quick on-line search for image analysis workshops will bring up thousands of internet pages.

PhosphorImagers

Digital images are acquired by direct video capture from wet gels using a charge-coupled device (CCD) camera, by scanning a previously generated photograph, or through the use of a PhosphorImager[®]. PhosphorImagers are laser-based detection devices with a wide linear dynamic range, allowing the observation and quantification of very weak bands and very strong bands, even when both are present on the same gel. The STORM[®] product line (Molecular Dynamics; GE Healthcare) was originally designed as a filmless autoradiography system for nucleic acid and protein band detection, and the phosphor detection technology reduced the amount of time necessary to produce an image. The later Typhoon[™] model (Molecular Dynamics; GE Healthcare) offers a much greater array of detection abilities across five orders of magnitude. In addition to imaging wet gels, PhosphorImager technology now supports autoradiography as well as detection by fluorescence and chemifluorescence, which is similar to chemiluminescence except for the final substrate. These instruments feature very powerful image analysis software and the ability to optimize, archive, and analyze data. For a more complete understanding of the PhosphorImaging detection approach, see Johnston *et al.* (1990).

Traditional methods of photodocumentation

In the traditional sense, photodocumentation refers to the investigator physically taking a picture of a gel soon after electrophoresis and staining. In the past, this was most commonly performed using Polaroid instant imaging products. The traditional systems for photodocumentation include the widely used handheld Polaroid DS-34 camera (Fig. 10.6) and the older-style Polaroid MP-4+ Multipurpose System. The DS-34 system in particular has proven to be an extremely versatile, low-cost instrument that remains a staple in many molecular biology labs. In either case the investigator takes a picture of a gel, tapes it into a laboratory notebook, and then annotates the picture appropriately. In contrast to the Polaroid instant imaging approach, newer instrumentation is firmly established in most laboratories. Digital technology systems for photodocumentation usually have an attached printer, thereby allowing the investigator to print a hard copy of the gel. Using these systems, the image is digitized and a hard copy is generated either on heat-sensitive thermal paper; otherwise, the image is simply sent to an on-line laser printer. Early on, thermal prints lacked definition, though improvements have been made. The high-gloss thermal paper currently in use in this laboratory produces particularly attractive gel images.



Figure 10.6 Polaroid DS-34 handheld camera. Snap-on hoods allow photodocumentation of a variety of gel sizes and they also position the camera for sharp focus. A built-in lens ensures proper magnification and focal length. Courtesy of Polaroid.

Sample visualization

Traditional photographs of gels are prepared either in black and white or in color, with or without a negative, using standard Polaroid film. In the case of nucleic acids, a gel is stained with SYBR Green, SYBR Gold, ethidium bromide (or other dye), UV irradiated from below by placing the gel on the surface of a standard laboratory transilluminator, and photographed or scanned. In some cases epi-illumination, i.e. UV irradiation from the above the gel, may enhance sensitivity. The distribution of the DNA or RNA in the gel is then recorded by simply clicking the shutter release on a Polaroid camera. If desired, images can then be characterized by digital analysis to determine mass, molecular weight, relative abundance of and ratios among bands on the gel. Digital systems are generally focused manually, and include a zoom feature to allow the investigator to photograph the entire gel or focus in on a region of interest. Polaroid systems accept a variety of instant film formats, and feature snap-on hoods that accommodate gels of several dimensions. These produce sharp electrophoretograms by maintaining the camera at the correct focal distance from the gel, which precludes the need to focus.

All image analysis systems allow the user to set the amount of time that the shutter in the lens will be open (the shutter speed) as well as the aperture setting (i.e., the *f*/stop, how wide the lens will open. If one is uncertain as to the correct parameters, try photographing ethidium bromide-stained gels for 0.5 s at an aperture of *f*/8 is a good start, and corrections can be made as needed. For SYBR Green- or SYBR Gold-stained gels, one should begin with 0.5 s at *f*/5.6. Of course, the exact settings will be empirically determined for each gel and are directly related to the mass of nucleic acid in the gel, the staining time, the destaining time (if any), the magnitude of background fluorescence, and whether epi-illumination or direct illumination from beneath the gel is used.

Filtration

Unless the bands stand out in a photograph of a gel, the images will not convey a useful level of detail. Since most electrophoresis photography involves black and white film or gray-scale image analysis, adequate contrast is essential. It should be noted, however, that color photographs of SYBR Green- and SYBR Gold-stained gels are stunning. Regardless of the camera or film used, color filters enhance detail and make electrophoresis bands appear more sharply defined.

A filter absorbs its own complementary colors and transmits its own color, allowing the photographer to manipulate the relative brightness of colors that would otherwise appear similar in black and white. For example, a red filter makes green and blue appear darker, and red appears relatively lighter. A useful mnemonic by which to remember pairs of complementary colors and, thereby, the correct filter for use, is as follows:

“Red Cadillac by General Motors”

that is

Red Cadillac: Red and Cyan
 by: Blue and Yellow
 General Motors: Green and Magenta

Table 10.2 lists some of the common gel electrophoresis stains² for nucleic acid and protein gels, the corresponding filters that offer the best results, and approximate aperture setting and shutter speeds. Because filters absorb light, it is necessary to ensure that the gel is properly illuminated; this means uniform irradiation from beneath. Burned out UV bulbs in transilluminators should be replaced by qualified personnel as soon as possible. To an extent, one may compensate for the effect of a filter by increasing the exposure time, using a wider aperture, or both. Moreover, because UV light sources have a low light output, high-speed films are particularly well suited for this purpose. In each application, the precise parameters are best determined empirically.

Tips for optimizing electrophoretograms

1. *Maximize the size of the image.* When using the DS-34 camera, be sure to select the hood that most closely approximates the size of the gel. Minigels (e.g. 7 cm × 8 cm, photographed with a large hood (15 cm × 15 cm) will cause the gel to look like a postage stamp in the middle of large, empty background. An image in which the

Table 10.2 Common Electrophoretic Stains and the Corresponding Color Filters

Light source (wavelength)	Stain	Filter	Aperture	Shutter speed ¹
UV (302 nm)	Ethidium Bromide	#22 Deep Orange or #15 Deep Yellow	f/8	1/4 second
UV (302 nm)	SYBR Green SYBR Gold	#15 Deep Yellow	f/5.6	1/2 second
Epi-illumination (254 nm)	SYBR Green SYBR Gold	#15 Deep Yellow	f/5.6	1/2 second
White (400–700 nm)	Coomassie Blue	#8 Medium Yellow or #9 Deep Yellow	f/16	1/30 second
White (400–700 nm)	ELISA	#8 Medium Yellow	f/16	1/30 second
White (400–700 nm)	Silver	#58 Green	f/16	1/30 second

¹Based on Polaroid Type 667 and 57 film (ISO 3000). Longer exposure times or wider apertures are necessary with lower-speed films. Filtration also applies to electronic video capture, though the need and effectiveness of integration must be determined empirically.

²For acridine orange staining and color photography, use Polaroid Type 108 color film and a yellow filter; for black and white photography use Polaroid Type 105 or Type 107 film with a red filter.

gel more or less fills the photograph will always offer the greatest level of detail, in terms of both sensitivity and resolution. For digital image analysis, zoom in on the bands of interest for maximum clarity.

2. *Use higher speed films with UV light.* Whenever possible, use films with ISO 3000, such as Polaroid Type 667 and Type 57 films, both of which are particularly well suited for most routine applications. High-speed films preserve depth of field and image sharpness, and shorter exposure time will minimize exposure of the investigator to UV light if using an older style transilluminator. If working with ethidium bromide, minimizing exposure of the gel to UV light will also facilitate taking multiple photographs because ethidium bromide is rapidly bleached by UV light to the extent that it will no longer fluoresce. Shorter exposure times will also minimize nicking of the phosphodiester backbone of nucleic acids because of over-irradiation of the gel. This is a very important consideration if a band is to be eluted from the gel for cloning and subsequent characterization. Although *minimal* exposure to UV light randomly nicks the backbone of RNA (and DNA), thereby enabling efficient Northern (or Southern) transfer, excessive nicking will render the sample incapable of nucleic acid hybridization in subsequent steps.
3. *Keep camera rollers clean.* To prevent spotting and other processing defects, clean DS-34 camera rollers periodically with warm water and a cotton swab or Kimwipe. The need to clean the rollers becomes obvious when three or four dots, evenly spaced, appear on the film (Fig. 10.7). These dots are the result of debris on the rollers, which places uneven pressure on the film as it is pulled from the camera.
4. *Keep transilluminator surface clean.* Gels should be removed from the surface of the transilluminator immediately after image capture. Gels that are allowed to remain on the transilluminator are often forgotten, only to be discovered days later dried out and stuck to the surface of the transilluminator. Also, the drying of ethidium bromide or other fluorescent dyes contributes to background spots in subsequent photos/image captures. It is wise to wipe the surface of the transilluminator with dH₂O or 70% ethanol and wipe it dry with a Kimwipe immediately after each use. Nothing takes the place of photographing a gel against a clean background.
5. *Follow film development instructions.* Pay close attention to the development time and temperature recommended for each type of instant film. Underdeveloping or overdeveloping film, a common error, can adversely affect the characteristics of

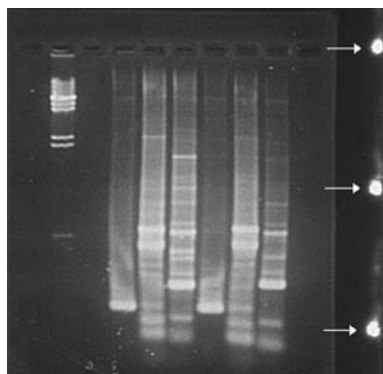


Figure 10.7 Effect of dirty rollers in a Polaroid camera. Notice the three heavy white dots, evenly spaced, along the extreme right-hand edge of the image.

the image. In general, the most commonly used black and white films require a 30 s development after the film is pulled out of the camera. Overdevelopment will enhance contrast, and this can be good or bad, depending on the extent of variation in mass among bands on the gel. By increasing the contrast, bands with the greatest mass will stand out even more (Fig. 10.8). However, bands in which the mass of sample is minimal may actually begin to blend into the background of the image, leading the viewer to believe that no such bands exist.

6. *Standardize digital camera settings.* For the sake of consistency, pay careful attention to the aperture setting, integration setting, and shutter speed. It is important that these settings be standardized in a laboratory in order to support meaningful comparison from one experiment to the next.
7. *Avoid temperature extremes.* The ideal development temperature for Polaroid films is room temperature, generally 21 to 24° (70 to 75°F). If the lab is much warmer or cooler, or if the film was stored refrigerated until just prior to use, an adjustment in development time or exposure time may be necessary. Refer to the film package insert for guidance.
8. *Use small apertures.* The smaller the aperture (the higher the f /stop), the greater the depth of field and the more sharply defined the resulting images. Ideally, one should begin with $f/16$ and optimize by incrementally decreasing the aperture, as much as practical. In this laboratory, ethidium bromide-stained gels are routinely photographed for 1/4 second at $f/8$. SYBR-stained gels are photographed at $f/5.6$ for 1/4 or 1/2 s, adjusted as needed.
9. *Minimize camera movement.* Gels photographed with slower films require long exposure times when stained with ethidium bromide and, in particular, SYBR Green or SYBR Gold. To minimize camera movement in this lab, a cable release has replaced the pistol-grip shutter release, which easily unscrews from the handheld DS-34 camera. It is also far more practical to use faster films (see Item 2). For image analysis workstations, camera movement is usually not a concern.
10. *Eliminate ambient light.* If using a handheld camera, be sure to rest the camera hood directly on the light box; this will prevent ambient light from reaching the film. It is usually not necessary to turn off the lights in the room with this type of camera. Be certain, though, to turn off the transilluminator immediately after making the exposure. If using an image analysis work station, in most cases the transilluminator will not switch on unless the safety door is closed and the lock engaged. This also serves to protect the user from the dangers of UV light at the same time as preventing ambient light from interfering with the production of the image.
11. *Use correct filtration.* Consider filtration a standard component of photodocumentation. Color filters can improve the contrast of electrophoretograms, as previously

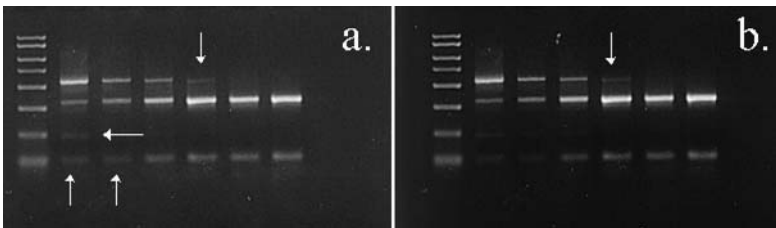


Figure 10.8 Effect of overdeveloping a Polaroid electrophoretogram. (a) Correct developing time. (b) Overdevelopment of film causes loss of low-intensity bands.

described. Ideally, one should use screw-in glass filters whenever possible. Acetate filters can sometimes distort the optical path and, for this reason, they should be avoided.

12. *Avoid fogging the lens.* Pay attention to the amount of time that the transilluminator has been turned on. Because the surface of the transilluminator becomes hot in a short time, some of the liquid from the gel will vaporize and then condense on the lens located on the inside of the hood or image analysis workstation. This effectively fogs the camera lens and compromises greatly the quality of the photograph. This is easily remedied by wiping the lens dry with a Kimwipe.
13. *Recover the film properly.* Film that has been exposed is to be pulled straight out of the camera, slowly and evenly, and perpendicular to the user. The film should not be pulled upward or downward as it is removed from the camera. It has been suggested that the user intone “Pol-ar-oid,” the cadence of which is used as a vocal guide to be followed as the film is pulled. When the recommended developing time has elapsed, turn the film-receiver sheet complex over, take hold of the corner of the *film*, where the number is printed, and peel it away from the receiver sheet, which is impregnated with the chemistries required for image development. Some films, including type 667, do not require any type of coating; the film may be slightly arched at first, but will flatten out within 4–5 min. Do not touch the negative itself or the sheet to which it is attached: both contain an alkali that should not come into contact with skin.

Finally, for the non-photographically inclined, the following should be taken to heart:

1. The shutter speed scale found on the lens is labeled “B, 1, 2, 4, 8, 16, 32, and 64”. The “B” setting means that the lens will be open as long as the shutter release is depressed by the user. The numbers that follow are the reciprocal of the actual shutter speed (i.e., 2 is actually 1/2 [0.5 s]; 4 is actually 1/4 [0.25 s], and so forth).
2. The terms “aperture setting” and “*f*/stop” are used interchangeably, and refer to how wide the lens shutter will open. Obviously, this setting influences the amount of light that will enter the camera and expose the film. The *f*/stop scale is also “anti-intuitive,” meaning that the smaller the number, the larger the opening of the shutter.

Inherent limitations of photographic and X-ray films

In photographic systems, the method of image recording is an energy conversion process whereby the sensitivity of silver halide crystals is used to record the level of photons emitted by or reflected from a source or object. The photons are absorbed by the silver crystal and when the level of energy is sufficient, the crystal forms a deformation known as a latent image site. At this site, the silver and halogen bond of the crystal is broken, resulting in free halogen and metallic silver. If the crystal is exposed to too much energy in too short a time, it will be unable to react fast enough to record the event “correctly”; likewise, if a low level of energy strikes the crystal for a prolonged time, then the crystal will react differently than it would under normal conditions.

At some point, the crystal will be exposed to sufficient energy to fully saturate it with metallic silver. Beyond this point, it can no longer record further information, becoming uniform in response at whatever level of response that system can produce. In the case of Polaroid film, a system referred to as direct positive response, this level of saturation relates to the lowest density, known as D-Min. When the crystal reaches such a level of low absorption of photon energy that it no longer is capable of producing a latent image, it acts as if no energy has been absorbed. This produces a uniform response of maximum density, known as D-Max. Between the D-Max and D-Min are regions where the silver crystal is recording the highlight and shadow details, in a nonlinear manner, as the crystal begins to approach the linear portion of its ability to record the energy absorption. The region of useful recording between the highlight and shadow region of the film is known as the dynamic range of the film. Although ideally the film and the photon source would have the same range, thus allowing the system to record in a true-to-life manner, it is often true that the photon source exceeds the range of the film. In this case, the user of the imaging system must choose which areas are of greatest importance. By adjusting the exposure of the system, the investigator defines the range of the image that will be recorded, while losing highlights and shadows outside the dynamic range of the film. In the molecular biology laboratory, where both strong and weak bands are often observed in the same image, it is laudable to make two or three exposures of the same image, using different aperture settings, in order to derive as much useful data from a single image as possible.

References and suggested reading

- Baxes, G.A. (1994). *Digital Image Processing – Principles and Applications*. John Wiley & Sons, New York, NY.
- Baxes, G.A. (1988). *Digital Image Processing – A Practical Primer*. Cascade Press, Denver, CO.
- Castleman, K.R. (1996). *Digital Image Processing*. Prentice-Hall, Inc, Englewood Cliffs, NJ.
- Chellappa, R. (1992). *Digital Image Processing*. IEEE Computer Society Press, Los Alamitos, CA.
- Gonzalez, R.C. and P. Wirtz. (1987). *Digital Image Processing*. Addison-Wesley Publishing, Reading, PA.
- Johnston, R.J., S.C. Pickett and D.L. Barker. (1990). Autoradiography using storage phosphor technology. *Electrophoresis* **11**, 355–360.
- Media Cybernetics (1996). In “Gel-Pro Analyzer Reference Guide,” 2nd edition. Media Cybernetics, Silver Spring, MD.
- Pratt, W.K. (1991). *Digital Image Processing*, 2nd edition. John Wiley & Sons, New York, NY.
- Russ, J.C. (1995). *The Image Processing Handbook*. CRC Press, Boca Raton, FL.

This page intentionally left blank

11 Northern Analysis

Rationale	241	Post-transfer handling of filters	254
Choice of filter membrane	242	Formaldehyde denaturing systems	254
Nitrocellulose	243	Glyoxal denaturing systems	254
Nylon	244	Option 1	254
Polyvinylidene difluoride	245	Option 2	255
Handling and filter preparation	245	Immobilization techniques	255
Northern transfer techniques	246	Baking	255
Capillary transfer	246	Crosslinking by UV irradiation	256
TurboBlotter™	247	Protocol: UV crosslinking RNA to nylon	
Vacuum blotting	248	filters	257
Electroblotting	249	Post-fixation handling of filters	257
Alkaline blotting	249	Reverse northern analysis	258
Protocol: RNA transfer by passive		References	259
capillary diffusion	250		
Protocol: TurboBlotter downward			
transfer of RNA	252		

Rationale

Many variations on the theme of Northern analysis have emerged since the original descriptions of the technique (Alwine *et al.*, 1977; 1979). At the heart of the Northern analysis is the transfer, or blotting, of electrophoretically separated RNAs from a gel to a filter membrane for subsequent fixation and hybridization to specific probes. Because the RNA is transferred in exactly the same configuration as it was separated in the gel, hybridization signals generated by Northern analysis provide a qualitative and semi-quantitative profile of the sample: one gains a snapshot of mRNA abundance and size. In addition, blotting the RNA from the gel onto a filter membrane has the effect of concentrating the RNA onto a relative small area on the filter, which helps with assay sensitivity. Northern analysis is compatible with RNA from tissue samples as well as cells grown in culture and continues to be a commonly used technique to assess mRNA length. Northern analysis, however, is not a very quantitative assay.

There are various methods for isolating RNA, preparing and running denaturing agarose gels, and transferring the nucleic acid from the gel. Selections must be made in advance by the investigator regarding the type of filter membrane,

method of immobilization, type of probe, method of probe labeling, hybridization recipe, and method of detection. Obviously, quite a few decisions must be made, and it is well to outline the process in advance so as to avoid surprises halfway through the assay. Optimization of Northern analysis comes from experience, and often involves “tweaking” at several points throughout the process. When optimized in one’s own laboratory, many of the technical difficulties and issues pertaining to reproducibility resolve themselves.

What then constitutes Northern analysis? Generically, it is the electrophoresis of RNA, followed by transfer and immobilization on a filter, hybridization with a labeled DNA or RNA probe, and finally detection and analysis of hybridization events. The terminology “Northern” analysis is something of a pun, since the original method for electrophoresis and blotting of DNA was coined “Southern” blotting in honor of the gentleman who developed it (Southern, 1975); it is a method used to study gene structure, gene organization, and gene copy number. Previous chapters have described methodologies for RNA isolation, quantification, and electrophoresis. This chapter examines the options regarding choice of filter membranes, transfer techniques, and methods for nucleic acid fixation on the filter. Subsequent chapters describe hybridization and detection methods.

Choice of filter membrane

One important decision associated with the study of RNA by Northern analysis is the selection of the filter membrane to which electrophoretically resolved RNA is transferred prior to hybridization. This selection must be influenced by the knowledge that different types of filters exhibit wide variations in nucleic acid binding capacity and support different transfer and immobilization methodologies. Consideration must also be given to the method of probe detection (autoradiography or chemiluminescence) because some membranes are quite incompatible with non-isotopic detection methods. Moreover, because the RNA sample on a filter paper may well represent labor-intensive efforts, many investigators perform repeated hybridizations using a battery of probe sequences. This requires that each hybridized probe be removed, following detection, by a very high-stringency wash; some types of nucleic acid blotting filters withstand the removal of previously-hybridized probe much better than other filter types. Suggestions for the proper handling of nucleic acid filter papers are described in detail by the manufacturer of the filter paper, and some general guidelines are described here.

Filter membranes for nucleic acid blotting can be purchased pre-cut to match the size of gels that are used in the lab. Alternatively, filters are often available as three-meter rolls, allowing the user the flexibility to custom cut a filter, as needed. In all cases, the filter membranes are sandwiched between two sheets of paper that protect the membrane from scratching and from particulate matter. The protective paper is usually printed with the name of the manufacturer and the trade name of the product while the filter itself is the unmarked white sheet

in the middle. It is very important to avoid touching the filter with ungloved fingers, as finger greases will prevent the proper wetting of the filter (described later), reduce the binding capacity of the membrane, and introduce RNase into the system. As needed, the filter paper can be picked up by the corner with nuclease-free forceps or gloved fingers. Laboratories utilizing nucleic acid blotting methods should invest in powder-free gloves, because glove powder can contribute to background smudges following hybridization detection. Filter membranes should be stored in a cool dry location, out of direct sunlight and should not be refrigerated, frozen, or heated above room temperature prior to first use. If stored correctly, filters are stable for many months.

Nitrocellulose

The original RNA blotting technique described the use of diazobenzoyloxymethyl paper for nucleic acid immobilization (Alwine *et al.*, 1977). This technique evolved fairly rapidly to accommodate the use of nitrocellulose, the so-called “classical membrane” of molecular biology. The underlying chemical basis by which nucleic acids become fixed onto nitrocellulose is hydrophobic in nature. Perhaps the most critical parameter is the need for very high ionic strength transfer buffer to promote the binding of single-stranded molecules (RNA or denatured DNA). The most common pore size for nitrocellulose used in blotting is 0.45 μm , although 0.22 μm membranes can also be used, particularly in the study of smaller nucleic acid molecules. Although nitrocellulose is still used today in some laboratories for nucleic acid analysis, there are intrinsic qualities that make the use of other membranes far more desirable:

1. Nitrocellulose has a relatively low binding capacity for nucleic acids (about 80–100 $\mu\text{g}/\text{cm}^2$).
2. Nitrocellulose has a very poor binding capacity for smaller nucleic acid molecules (typically less than 1000 bases). This precludes efficient blotting of most PCR products.
3. Nitrocellulose has a slight net negative charge that mandates the use of a relatively high ionic strength transfer buffer, usually 20 \times SSC or 20 \times SSPE¹.
4. Nitrocellulose requires baking in a vacuum oven at 80° to immobilize nucleic acids. Baking nitrocellulose in a vacuum oven is a non-negotiable requirement because humidity will interfere with the immobilization process.
5. Nitrocellulose becomes very brittle after the baking step that is required for nucleic acid immobilization. This makes subsequent manipulations very difficult and requires handling of the membrane with extreme care.
6. The frailty of this membrane following the baking step makes repeated probing impractical.
7. The background associated with the use of nitrocellulose for non-isotopic methods can be excessive.

¹SSC (saline-sodium citrate): 20 \times SSC = 3M NaCl; 0.3M Na₃-citrate; adjust to pH7.0. SSPE (saline-sodium phosphate-EDTA): 20 \times = 3M NaCl; 0.2M NaH₂PO₄·H₂O; 0.02M EDTA; adjust to pH7.4.

The rather flimsy nature of nitrocellulose was the impetus behind the development of reinforced nitrocellulose filters, an example of which is Optitran™ (Whatman). This membrane retains all of the properties of nitrocellulose and has a polymer backing; it is therefore much stronger than ordinary nitrocellulose and thus able to withstand more aggressive manipulation.

Nylon

Nylon filters were developed to circumvent many of the difficulties inherent in nitrocellulose and have largely replaced nitrocellulose for most routine nucleic acid blotting applications. Perhaps the most significant advantage associated with their use is that these membranes exhibit great tensile strength. This enables very assertive handling of the filter and repeated probing of the same filter. Nylon filters are widely used in the molecular biology laboratory and are sold under a variety of trade names. They fall into two categories: those with a neutral surface charge, an ancient name for which is nylon 66, and those with a net positive charge imparted by surface amines. The latter are frequently referred to as charge-modified nylon filters and denoted as nylon(+), which are available with varying charge densities, depending on the manufacturer. Nylon filters used for blotting are available in 0.2 μm and 0.45 μm pore sizes, as is nitrocellulose. Nylon filters, particularly the nylon(+) variety, have also proven to be very compatible with chemiluminescence-related technologies. Nylon membranes exhibit an enhanced nucleic acid binding capacity compared to nitrocellulose, in some cases as much as 400–500 $\mu\text{g}/\text{cm}^2$, depending on the manufacturer and the exact buffers used by the investigator. Since nylon membranes also show a particular affinity for smaller nucleic acid molecules (500 bases), routine blotting of small molecules such as PCR products can be performed readily. Compared to nitrocellulose, the transfer buffers used in conjunction with nylon filters are of lesser ionic strength, often as low as $5\times$ SSC or $5\times$ SSPE. While the positive charge associated with nylon(+) membranes certainly increases the electrostatic attraction between target RNA and membrane, adequate blocking of the membrane during prehybridization and, if applicable, during the detection process normally produces blots with exceptionally clean background. However, failure to prehybridize or block properly will most certainly produce excessive background due to non-specific interaction between the positively charged membrane and the negatively charged phosphodiester backbone of the probe.

The investigator is afforded much greater latitude when blotting onto nylon membranes. Immobilization of nucleic acids can be accomplished either by baking or by irradiation with UV light. The baking of nylon filters does not require a vacuum oven and can be performed in any instrument capable of generating 65 to 70° for 1.0 to 1.5 h. Best of all, nylon filters do not become brittle when baked, and they withstand repeated probing much better than nitrocellulose. Finally, it has been reported that torn nylon filters can be repaired by spot-welding them with a hot metal implement such as a bacterial inoculating loop

or soldering pen (Pitas, 1989). In this laboratory, partially torn filters generally do not contribute to background; by wrapping the filter in plastic wrap, the adjoining sides of the tear in the filter are usually not noticeable on the X-ray film following detection.

Polyvinylidene difluoride

Polyvinylidene difluoride (PVDF) is a fluorocarbon polymer that historically has been a favorite membrane for western blotting. The use of this membrane for Northern and Southern analysis has been described in the literature and used successfully in this lab, in conjunction with non-isotopic probes, by handling the PVDF filters as if they were nylon membranes² PVDF is a high durability, chemically stable membrane that is sold under a number of different trade names. These membranes possess an average pore size of 0.45 μm . Nucleic acids can be permanently immobilized by crosslinking with UV light or by baking (vacuum oven not required).

Handling and filter preparation

No matter what type of filter membrane is selected, it should be handled only with gloves, as described above, and cut with sharp scissors or a scalpel. When cutting the membrane to the appropriate size, which will usually be the size of the gel, the two sheets of protective wrapping in which the filter is provided should be left in place and used to handle the filter. Unused filter membranes should be returned to storage right away in order to prevent the accidental spillage of anything onto the filter or accidental contamination with RNase or DNase.

Filter membranes generally require pre-wetting in a clean pan or tray of sterile water for 2 min, followed by brief equilibration in the high-salt transfer buffer selected for the particular application at hand. In this laboratory, a special supply of trays is reserved exclusively for use in the preparation of filters for Northern transfer. In addition to handling only with gloves, these trays are also rinsed in H_2O_2 for several minutes and then rinsed with nuclease-free H_2O prior to use. The proper way to wet a filter is to float it on the surface of nuclease-free water. An immediate color change in the appearance of the membrane from white to gray should be observed as the filter comes into contact with the water. This technique is strongly recommended because it will allow immediate identification of any area of the filter that has become mechanically damaged, touched with bare fingers, stored incorrectly, or otherwise compromised: such areas appear as smudges or an obvious uneven wetting of the filter.

²When used for Northern analysis, PVDF membranes (e.g. Immobilon-P, from Millipore) were pre-wet in methanol for 15 s (per the directions of the manufacturer), wet for a minimum of 2 min in water, and then for 30 s in $5\times$ SSC. The membrane was then placed in contact with the gel and the transfer apparatus assembled.

Once it is established that the filter has not been damaged, it can be completely submerged by rocking the tray. Allow the filter to pre-wet for at least 1–2 min (longer periods are fine, too).

After wetting, the filter is transferred to a tray containing the same transfer buffer that will be used for the blot. A brief equilibration of about 30 s is all that is required, though the filter can rest in transfer buffer until ready to use. Note that the order of wetting is much more important than the time: water first, followed by equilibration in high-salt transfer buffer. The filter is then ready for immediate use in the transfer process itself. One may mark the filter asymmetrically now or, preferably, wait until the transfer is complete, as described below.

Northern transfer techniques

The actual transfer of electrophoretically separated RNA species from the gel to a solid support can be accomplished using any of a variety of blotting formats; these differ primarily in the method by which the RNA is drawn out of the gel matrix. No matter which transfer technique is chosen, *intact* RNA must have been electrophoresed in a denaturing gel (for maximum resolution) under conditions that maintain RNA integrity (for maximum sensitivity). Otherwise, the resultant data are likely to be suboptimal. In general, the transfer techniques described here are applicable to both Northern and Southern analysis.

Capillary transfer

The method requiring the fewest accessories is the traditional passive capillary transfer method, which has also been used in the classical Southern analysis (Southern, 1975). It is also the least expensive, and most time consuming, technique. In short, a stack of absorbent material, such as paper towels or blotting paper, is used to draw (wick) the transfer buffer from a reservoir, through the gel, and finally into the dry stack of paper towels. In so doing, the RNA sample is essentially eluted from the gel and is trapped on a juxtaposed filter membrane, as the buffer continues to move through the membrane and into the stack of paper towels. Traditionally, a typical capillary transfer setup involves placing the blotting paper on the top of the stack, as shown in Fig. 11.1. Two common complaints associated with this transfer method, however, are (1) poor transfer efficiency, especially of larger RNA molecules; and (2) transfer takes too long.

Regarding the former, inefficient transfer is commonly associated with the following:

1. The gel was too thick.
2. The gel consisted of more than 1.2% agarose.
3. EtBr was used to stain the gel, either during electrophoresis or afterward, without proper destaining (recall that there are excellent alternatives to staining with EtBr).
4. The gel was not soaked long enough to remove the formaldehyde prior to setting up the transfer.

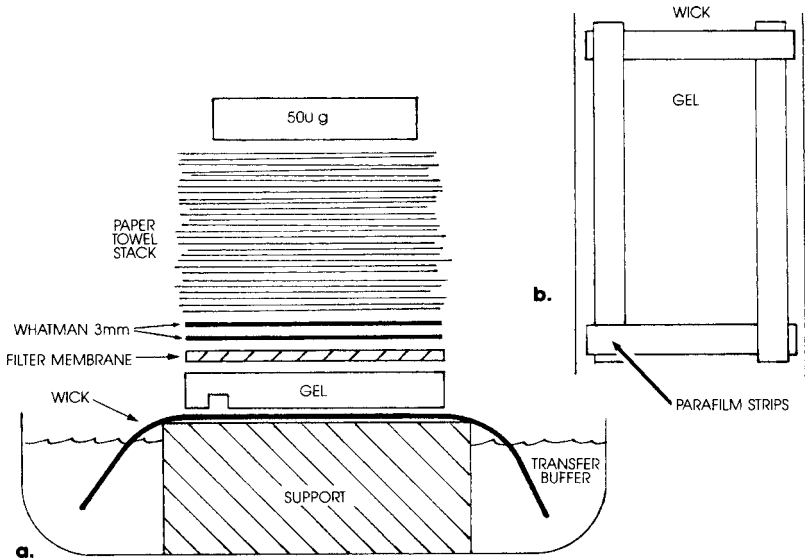


Figure 11.1 A typical setup for capillary transfer of RNA from denaturing agarose gels to a filter membrane. (a) Side view. Invert the gel to reduce the blotting time and achieve greater resolution, since the RNA does not have to travel as far through the gel matrix to reach the membrane. (b) Top view. Place overlapping strips of Parafilm directly on the wick and up against the gel (not on top of the gel). Set Parafilm in place around the sides of the gel first, and then overlay filter membrane and other blotting materials.

5. The gel was not UV irradiated to nick the backbone and promote expedient transfer.
6. The transfer period was too brief. Most capillary transfers require 16 h or longer, i.e., overnight.
7. The RNA was degraded to begin with, so no hybridization signal could possibly be generated, no matter how efficient the transfer.
8. Any combination of the previously listed items.

In instances where incomplete transfer simply cannot be remedied and, regarding the latter complaint of time-consuming transfer, alternative transfer techniques that accelerate and improve transfer efficiency have been developed and are in widespread use, some of which are described below.

TurboBlotter™

TurboBlotter (Whatman) downward transfer systems take advantage of gravity and, in so doing, offer enhanced transfer efficiency in a much shorter time, usually less than 3 h, than is required by the older upward capillary transfer method (Fig. 11.2). It has also been shown empirically in this lab that, when using the Turboblotter apparatus, most of the nucleic acid transfers out of the gel in the first hour. Pre-cut filters and blotting papers are available to accommodate all

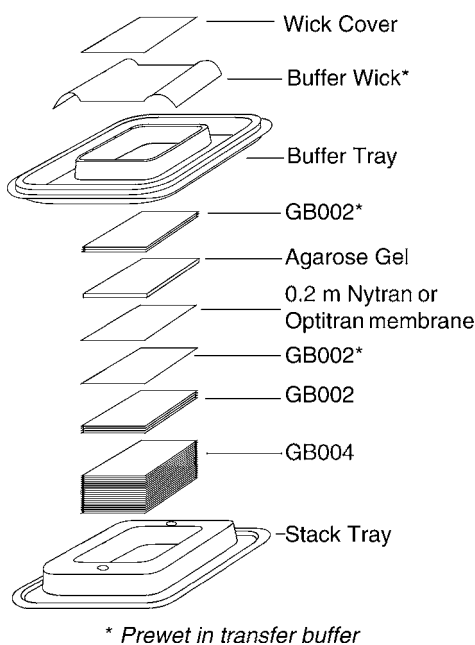


Figure 11.2 TurboBlotter transfer method. The principle is essentially that of an inverted capillary transfer. By taking advantage of gravity, rapid downward movement of the high-salt transfer buffer accelerates the transfer and dramatically improves transfer efficiency. The absorbent blotting paper is at the bottom of the stack and the buffer reservoir on top, precisely the opposite of the classical capillary blot setup shown in Figure 11.1. Courtesy of Whatman, Inc.

standard-sized gels and the TurboBlotter device has a very small footprint on the lab bench. This method likewise eliminates the need for heavy weights on top of the capillary stack, thereby minimizing compression and concomitant collapsing of the pores of the gel. The technique does not require the use of any expensive equipment, such as the vacuum blot devices or electroblotting apparatuses described later, and is extremely simple to assemble. The record for assembling a TurboBlot transfer in this lab is 58 s!

Vacuum blotting

In a vacuum blotting system, negative pressure is applied to accelerate the transfer process. This technique is also known as vacuum-assisted capillary transfer, in which the transfer buffer is essentially sucked out of the gel along with the nucleic acid sample. There are, of course, a number of different instruments for vacuum blotting, available from different vendors. The advantage of vacuum blotting is that the transfer occurs far more rapidly than in the traditional capillary transfer (30 to 60 min vs. overnight). To an extent, the accelerated process reduces molecular diffusion of nucleic acid molecules on the way out

of the gel, minimizing the outward (lateral) spreading of otherwise very sharp bands observed in the gel. One must use great care when vacuum-blotting, however, not to exceed the hydrostatic pressure of the gel (vacuums between 30 to 50 mbar are optimal), because gel collapse will result in immediate ruination of the experiment. Vacuum blotting is used most often for the rapid transfer of nucleic acids from agarose gels (Medveczky *et al.*, 1987; Olszewska and Jones, 1988); vacuum blotting of nucleic acids from polyacrylamide gels is also possible, though electroblotting (below) will yield superior results.

Electroblotting

A completely different approach known as electroblotting is yet another transfer method. Electroblotting is widely used to transfer proteins from polyacrylamide gels, and is less frequently utilized for the transfer of nucleic acids from agarose gels. One form of this technique known as electrophoretic transfer is performed by placing the gel next to the membrane in a special cassette that, in turn, is placed in a tank of electrolyte buffer (tank electroblotting). Upon application of a voltage gradient perpendicular to the direction of the initial electrophoresis the gel, the sample migrates out of the gel and onto the filter paper; this is essentially the standard transfer methodology for western blots. Another form of this technique, known as semi-dry electroblotting, requires only a minimal amount of buffer to saturate the membrane and blotting papers, establishing electrical contact between the gel and apparatus itself. The advantage here is that low voltage and current are needed for transfer, precluding the requirement for a high-current power supply. A wide selection of electroblotting apparatuses is commercially available (e.g., Hoefer[®] product line). The extent to which electroblotting accelerates the transfer procedure is directly related to the current; the investigator should follow the instructions of the manufacturer and exercise due caution when working with power supplies and other high-voltage equipment.

Either nylon or nitrocellulose may be used for electroblotting; the investigator should follow the recommendations of the manufacturer. Electroblotting can also produce a significant amount of heat, and many protocols strongly recommend cooling the electrolyte buffer prior to use as well as controlling the temperature during the run. For an historical perspective see Arnheim and Southern (1977) and Bittner *et al.* (1980).

Alkaline blotting

Alkaline blotting of nucleic acids is essentially a variation on the theme of capillary transfer. The key difference between these two approaches, however, is the method of denaturation and nucleic acid transfer. Traditional approaches to capillary blotting include the use of high-salt buffer (15–20 × SSC or SSPE) to blot RNA or previously alkali-denatured DNA from the gel to the filter membrane (positively charged nylon is recommended most often). In alkaline blotting of DNA, the transfer steps are accomplished simultaneously by the use of

0.4N NaOH (Chomczynski and Qasba, 1984; Reed and Mann, 1985), with subsequent fixation onto nylon membranes. Subsequently, it was shown that RNA can be transferred in 8 mM NaOH (Chomczynski, 1992) for no more than 1.5 h; the limited alkaline hydrolysis which ensues can facilitate the transfer of RNA from the gel to a solid support. A variant of this method is applicable to the blotting of double-stranded RNA as well (Li *et al.*, 1987). Succinctly, one must take care to perform RNA alkaline blotting correctly. Depending on the exact parameters used, alkaline blotting onto certain membranes can preclude the requirement for post-transfer immobilization. Moreover, the exact composition of transfer buffer and transfer time greatly influence the signal-to-noise ratio, with regard to the non-specific binding of nucleic acid probes to rRNAs, particularly when the alkaline transfer blotting method is used. This approach is also compatible with non-isotopic probes (L6w and Rausch, 1994).

Protocol: RNA transfer by passive capillary diffusion

Note: Refer to Fig. 11.1 as needed.

1. Wear gloves throughout the procedure. It is imperative to prevent nuclease contamination at every level. Remember that RNAs are not safe from nuclease degradation until they have been immobilized, by baking or crosslinking, onto a filter membrane.
2. Gels that have been stained with ethidium bromide or that contain formaldehyde should be soaked in two to three changes of nuclease-free water or buffer prior to setting up the transfer. Failure to do so will result in highly inefficient transfer, loss of sample, and impaired hybridization efficiency.

Note: A wide spatula is most convenient for transferring gels from one solution to another. Formaldehyde gels are especially slippery and should be handled with care.

- 2a. Formaldehyde gels: soak the gel in an excess of $5\times$ SSC or nuclease-free H_2O to remove the formaldehyde. Formaldehyde is typically drawn out of unstained gels by soaking them three times for 10 min each. If the gel also contains ethidium bromide, it will likewise be drawn out of the gel. This is required to eliminate the high background fluorescence associated with EtBr-stained formaldehyde gels so they can be photographed. Thorough destaining, while not essential, may require two or more hours. In this laboratory, formaldehyde gels are occasionally destained overnight in an excess of buffer at 4° . A gel stained in SYBR Green or SYBR Gold does not require destaining, though soaking the gel in H_2O or $5\times$ SSC will facilitate the removal of formaldehyde.

Optional, although not recommended: soak the gel in 50 mM NaOH, 10 mM NaCl for 30 min; in 100 mM Tris, pH 7.5 for 30 min, and then in $10\times$ SSC for 30 min. This will result in partial alkaline hydrolysis along the length of the RNA backbone. While this may facilitate transfer of very large RNA molecules out of the gel it is usually not necessary. Moreover, over-exposure to NaOH can destroy the ability of the RNA to hybridize to any probe. Instead, random nicking of the RNA backbone and improved transfer efficiency may be accomplished by irradiating the gel on a UV transilluminator or, for greater precision, by using a calibrated UV light source.

- 2b. Glyoxal gels: Transfer is set up immediately after staining (destain as needed) and photographing the gel. Glyoxal will be removed from the RNA in a post-transfer, prehybridization wash.

3. Trim the left and right edges of the gel (but not the top or bottom) such that there is a 1 cm margin on each side bordering the lanes of interest.
 4. Cut a piece of filter membrane to the size of the gel. Never handle the filter without gloves. Finger greases will compromise the binding capacity of most filters. To cut the filter membrane and other sheets of blotting paper to the size of the gel, place a clean, dry, gel casting tray directly on the wrapping paper that protects the membrane itself, and trace the outline of the casting tray with a pencil. The gel casting tray acts as a template so that all materials (filter, Whatman paper, paper towels) can be cut to the precise size of the gel without a ruler. This can be done while the gel is destaining or soaking in buffer. For laboratories with enhanced budgets, all required blotting materials may be purchased pre-cut to the proper dimensions.
 5. Prepare the filter for transfer according to the manufacture's instructions. Typically, nylon and nitrocellulose membranes are first floated on the surface of RNase-free water and then submerged for about 2 min. Old and/or soiled filters do not wet evenly, showing irregular patches of wetting when floated on the surface of water, and should be discarded.
 6. After wetting the filter paper for 2 min, the filter is next equilibrated in transfer buffer ($5\times$ SSC or $5\times$ SSPE for nylon) for 30 s or longer. The filter may remain in transfer buffer until it is used.
 7. Cut one sheet of Whatman 3MM paper (or the equivalent) approximately 2 cm wider than the gel and about 15 cm longer than the gel. This is the wick and it will act as the contact between the transfer buffer reservoir and the gel itself; a wick that is cut too long can be trimmed later.
 8. Saturate the wick by soaking it in transfer buffer for 10 s. Next, drape the wick over a plexiglass sheet supported by a baking dish or similar implement.
 9. Fill the baking dish with enough transfer buffer such that a minimum of 3 to 4 cm of each end of the wick is submerged in the transfer buffer.
 10. Any air bubbles that are trapped under the wick are easily removed by gentle rolling with a sterile pipette. Failure to remove all air bubbles will compromise transfer efficiency at that location.
 11. Place the gel in the middle of the wick with the wells facing down. Remove any air bubbles trapped between the gel and the wick. Mask the area surrounding the gel with strips of Parafilm to prevent short-circuiting³ the system during transfer (Fig. 11.1b).
 12. Carefully place the filter membrane on top of the gel and remove any air bubbles trapped beneath it.
 13. Cut two or three sheets of Whatman 3MM paper (or the equivalent) to the same size as the gel. Pre-wet at least one of these sheets in the transfer buffer.
- Note: Pre-wet at least one of these sheets in transfer buffer just prior to use. In this lab, pre-wetting the sheet that will rest directly on the filter membrane appears to prime the wicking action efficiently.*
14. Place the wet sheets of Whatman 3MM paper on top of the filter membrane and eliminate air bubbles. Then place the dry piece on top.

³Occasionally, the paper towels that draw the buffer upward from the reservoir and through the gel will sag down below the gel as they become wet. This may occur when the paper towels are merely slightly larger than the gel. If the paper towels touch the wick along the side of the gel, they will draw the transfer buffer up *around* the gel rather than through it. Less buffer moving through the gel slows the transfer process. To preclude this potential difficulty, some investigators have been known to cut away the wells from the rest of the gel (not necessary). Parafilm along the periphery of the gel will prevent direct contact between wick and paper towels.

15. Cut a stack of paper towels (5–6 cm when compressed) to the size of the gel. This can be done well ahead of time. Position the paper towels on top of the sheets of Whatman 3MM paper. Finally, cover the entire stack with a 500 g weight or an obscure textbook to compress the entire setup.

Note: Avoid paper towels with creases such as the C-fold type as inefficient transfer from the gel over the area of the crease is almost guaranteed. If no other paper towels are available, alternate the paper towels in the stack so that the creases are not aligned over any particular part of the gel. Uniformity of capillary action is an absolute requirement. As an alternative to paper towels, use pre-cut blotting paper.

16. Allow the transfer to proceed for several hours or overnight, if possible.
17. At the conclusion of the transfer period remove and discard the saturated paper towels. Then, carefully peel the filter membrane away from the gel with forceps or a gloved finger. Clip one corner of the filter to mark it asymmetrically and indicate orientation. This will be essential for accurate post-hybridization interpretation of data.
18. Follow the post-transfer washes suggested by the manufacturer. Typically, wash the membrane for 30 s in a fresh aliquot of the same buffer that was used for the transfer (e.g. $5\times$ SSC or $5\times$ SSPE).

Note: The brief, post-transfer washing of the filter removes any pieces of agarose that may be clinging to the filter as well as any other residual materials that could interfere with immobilization of the sample, hybridization specificity, or contribute to background.

19. Nylon membranes should be UV-irradiated while still damp but not sopping wet. Nitrocellulose membranes should be air-dried and then baked as soon as possible.

Protocol: TurboBlotter downward transfer of RNA

Note: Refer to Fig. 11.2 as needed.

1. Wear gloves throughout the procedure. It is imperative to prevent nuclease contamination at every level. Remember that RNAs are not safe from nuclease degradation until they have been immobilized, by baking or crosslinking, on a filter membrane (nitrocellulose, nylon, etc).
2. Gels that have been stained with EtBr or that contain formaldehyde should be soaked in two to three changes of sterile water prior to setting up the transfer. Failure to do so will result in highly inefficient transfer and loss of sample.

Note: A wide spatula is most convenient for transferring gels from one solution to another. Formaldehyde gels are especially slippery and should be handled with care.

- 2a. Formaldehyde gels: Soak the gel in an excess of $5\times$ SSC or nuclease-free H_2O to remove the formaldehyde. Formaldehyde is typically drawn out of unstained gels by soaking them three times for 10 min each. If the gel also contains ethidium bromide, it will likewise be drawn out of the gel. This is required to eliminate the high background fluorescence associated with EtBr-stained formaldehyde gels so they can be photographed. Thorough destaining, while not essential, may require two or more hours. In this laboratory, formaldehyde gels are occasionally destained overnight in an excess of buffer at 4° . Gels stained in SYBR Green or SYBR Gold do not require destaining, though soaking in H_2O or $5\times$ SSC will facilitate removal of formaldehyde.

Optional, though not recommended: soak the gel in 50 mM NaOH, 10 mM NaCl for 30 min; in 100 mM Tris, pH 7.5 for 30 min, and then in $10\times$ SSC for 30 min. This will result in partial alkaline hydrolysis along the length of the RNA backbone.

While this may facilitate transfer of very large RNA molecules out of the gel it is usually not necessary. Moreover, over-exposure to NaOH can destroy the ability of the RNA to hybridize to any probe. Instead, random nicking of the RNA backbone and improved transfer efficiency may be accomplished by irradiating the gel on a UV transilluminator or, for greater precision, by using a calibrated UV light source.

- 2b. Glyoxal gels: Transfer is set up immediately after electrophoresis has been completed. Glyoxal will be removed from the RNA in a post-transfer, prehybridization wash.
3. Select pre-cut blotting papers that are as close to the size of the gel as possible. If the gel is slightly larger, the excess will be trimmed away during the assembly of the TurboBlotter. If the dimensions of the gel are less than those of the blotting paper, the exposed portion of the blotting paper will be insulated later with Parafilm.
4. Cut a piece of nylon filter to the size of the gel. Never handle the filter without gloves. Finger greases will compromise the binding capacity of most filters. The other blotting papers need to set up this type of transfer are provided along with the TurboBlotter apparatus.
5. Prepare the filter for transfer according to the manufacturer's instructions. Typically, nylon (and nitrocellulose) membranes are first floated on the surface of RNase-free water and then submerged for about 2 min. Old and/or soiled filters do not wet evenly, showing irregular patches of wetting when floated on the surface of water, and should be discarded.
6. After wetting the filter paper for at least 2 min, the filter is next equilibrated in transfer buffer ($5\times$ SSC or $5\times$ SSPE for nylon) for 30 s or longer. The filter may remain in transfer buffer until it is used.
7. Place the stack tray of the transfer device on the bench, making sure that it is level.
8. Place 20 sheets of dry GB004 blotting paper in the stack tray.
9. Place 4 sheets of dry 3MM blotting paper on top of the GB004 stack.
10. Place 1 sheet of 3MM blotting paper, prewet in transfer buffer, onto the stack. Use a sterile 5 ml pipette to gently smooth the membrane, ensuring that no air bubbles are trapped beneath. Do not apply too much pressure with the pipette.
11. Place the equilibrated filter membrane on the stack.
12. Carefully place the agarose gel on top of the filter membrane and then trim the gel to the size of the membrane by cutting away any overhanging, unused lanes. *Do not* move the gel once it has made contact with the filter. If any portion of the stack is exposed because the gel was smaller, mask the area surrounding the gel with strips of Parafilm to prevent short-circuiting the system during transfer. This is usually not necessary with the TurboBlotter setup unless the gel is significantly smaller than the pre-cut blotting paper.
13. Use a sterile 5 ml pipette to gently smooth the gel, ensuring that no air bubbles are trapped between the gel and the filter paper.
14. Wet 3 sheets of 3MM blotting paper in transfer buffer, and carefully place these on top of the stack.
15. Use a sterile 5 ml pipette to gently smooth the stack, ensuring that no air bubbles are trapped between the gel and the sheets of blotting paper.
16. Attach the TurboBlotter buffer tray to the stack tray, guided by the circular alignment buttons to ensure proper placement.
17. Fill the upper buffer tray with transfer buffer. Do not pour any transfer buffer into the center of the stack tray where the blotting papers and gel are located.
18. Wet the wick in transfer buffer. Initiate the transfer by connecting the gel stack with the buffer tray: Place the wick across the stack so that the shorter dimension of

the wick completely covers the blotting stack and both ends of the long dimension extend into the buffer tray.

19. Cover the stack with the wick cover (provided) or simply with Parafilm. This is strongly recommended to prevent evaporation and salt precipitation. No weight is necessary on the stack.
20. Allow transfer to proceed until the buffer has been drawn out of the buffer tray. While this usually requires 3–4 h, it has been determined in this laboratory that for routine blots most of the sample in the gel transfers within the first hour (data not shown).
21. At the conclusion of the transfer period remove and discard the saturated wick and blotting papers above the gel. Then, carefully peel the filter membrane away from the gel with forceps or a gloved finger. Clip one corner of the filter in order to mark it asymmetrically to indicate orientation. This will be essential for accurate post-hybridization interpretation of data.
22. Follow the post-transfer washes suggested by the manufacturer. Typically, wash the membrane for 30 s in a fresh aliquot of the same buffer that was used for the transfer (e.g. $5\times$ SSC or $5\times$ SSPE).

Note: The brief, post-transfer washing of the filter removes any pieces of agarose that may be clinging to the filter as well as any other residual materials that could interfere with immobilization of the sample, hybridization specificity, or contribute to background.

23. Nylon membranes should be UV-irradiated while still damp but not sopping wet. Nitrocellulose membranes should be air-dried and then baked as soon as possible.

Post-transfer handling of filters

Formaldehyde denaturing systems

Although filters can be stored confidently before hybridization, it is necessary to immobilize the sample on the filter as soon as possible after transfer. Only then is the RNA sample stabilized and rendered resistant to degradation by RNase. Once immobilization is complete, filters can be used immediately for hybridization or stored for months. Place air-dried filters between sheets of Whatman 3MM paper and store dry and in the dark until further use. Formaldehyde will have already been removed from the system during gel destaining in preparation for gel blotting.

Glyoxal denaturing systems

The glyoxal that denatured the RNA prior to electrophoresis must now be removed from the filter.

Option 1

Air-dry and then bake the filter membrane according to the recommendations of the manufacturer. This will destroy some of the glyoxal and immobilize the

RNA at the same time. The remaining glyoxal can be removed by immersing the filter in 200 ml of 20 mM Tris, pH 8.0, preheated to 90°, and then immediately cooling to room temperature.

Option 2

After immobilization, remove glyoxal from RNA by washing the filter for 15 min at 65° in 20 mM Tris, pH 8.0.

Immobilization techniques

Thorough immobilization of target RNA and DNA on a filter membrane after the transfer has been completed is just as important as ensuring that the transfer process itself is complete. It is essential that the investigator understands that nucleic acids, especially RNA molecules, are not safe from nuclease or natural biological degradation until the sample has been immobilized onto a solid support. Failure to completely immobilize the sample will invariably result in the loss of the sample during the prehybridization, hybridization and/or post-hybridization washing and detection steps. A variety of methods for nucleic acid immobilization have been described, each of which has defined advantages, limitations, and applicability. In this laboratory, nylon filters are used almost exclusively because of the great experimental flexibility that these filters offer. In every instance, it is strongly advisable to follow the recommendations of the manufacturer of a particular filter membrane. A few general guidelines are presented here.

Baking

The classical technique for nucleic acid immobilization involves baking the sample onto the surface of nitrocellulose in a vacuum oven. Although the exact nature of interaction between the matrix and nucleic acids is not clear, it is believed to be hydrophobic. This method, in spite of its numerous shortcomings, persists today. The investigator allows the nitrocellulose filter to air-dry completely, after the post-transfer wash, and then bakes the membrane for 2 h at 80° *in vacuo*. When working with nitrocellulose, the use of a vacuum oven is a non-negotiable requirement because humidity interferes with the immobilization process. Whereas baking will certainly do the job, this method of immobilization is not a completely permanent process: after a few reprobings, some of the sample is lost from the surface of the filter, thereby compromising the quantitiveness of all subsequent assays.

With nylon membranes, nucleic acid samples can be immobilized by baking or by UV irradiation. The baking process usually requires 1 to 1.5 h, with temperatures only as high as 65°. Moreover, baking of nylon does not require a vacuum oven. Some protocols suggest that nylon filters may be baked while

still damp, although in this case, the application of a vacuum may accelerate the drying process and improve retention on the filter.

Crosslinking by UV irradiation

A more reliable method for immobilization is the use of a calibrated UV light source that can effectively crosslink nucleic acids⁴ to nylon membranes (Church and Gilbert, 1984; Khandjian, 1987); the practice of UV irradiation of nitrocellulose is to be avoided because it does not result in the comparable retention of nucleic acids achievable by crosslinking to nylon and, more importantly, over-irradiation of nitrocellulose poses a serious risk of fire. Don't do it.

Crosslinking of nucleic acids to filters with UV light has been a viable option since it was described as part of the Church and Gilbert (1984) protocol for genomic sequencing. Short-wave and medium-wave UV light (254 and 302 nm, respectively) activate thymine and uracil, and other bases to a lesser degree, which become highly reactive and form covalent bonds with the surface amines that characterize many nylon matrices (Saito *et al.*, 1981). UV crosslinking is usually more efficient with damp nylon membranes, which generally require a total exposure of 120,000 $\mu\text{J}/\text{cm}^2$. Alternatively, air-dried membranes require approximately 160,000 $\mu\text{J}/\text{cm}^2$.

Increased stability of nucleic acid on the matrix, compared to baking procedures, is an important dividend associated with this method of fixation. Many laboratories have reported an increase in sensitivity with UV-crosslinked filters compared to baked filters. Moreover, fixation by crosslinking appears to be a far more permanent immobilization technique than baking. This is of critical importance if the investigator plans to screen a filter more than once; repeated high-stringency removal of hybridized probe usually results in the loss of target sequence from filters that were baked.

A variety of methods have been derived to accomplish fixation. The most precise energy output is, of course, delivered by a calibrated UV light source (e.g., Stratalink[®], Stratagene, La Jolla, CA). Such instruments have an auto-crosslink feature preset to deliver a UV dose of about 120,000 $\mu\text{J}/\text{cm}^2$, ideal for a damp membrane, or any other amount of energy that may be desired. Alternatively, a standard laboratory transilluminator may suffice, though this will involve considerable trial and error in the absence of a calibration instrument. Most transilluminators in the molecular biology laboratory emit at one of two wavelengths, either 302 nm (medium-wave energy) or 254 nm (short-wave energy). Crosslinking with a 302 nm emitter is strongly recommended over 254 nm-emitting instruments because over-irradiation with the shorter wave energy will reduce sensitivity due to gross damage to the blotted nucleic

⁴In addition to a crosslinking function, irradiation of agarose gels is an alternative method of randomly nicking the phosphodiester backbone of RNA and DNA. This takes the place of limited hydrolysis by immersion in a dilute solution of NaOH, and is more controllable and, therefore, reproducible. Random nicking will facilitate nucleic acid transfer from the gel onto the membrane.

acid molecules. Damp filters are placed directly on the surface of the transilluminator and irradiated.

Protocol: UV crosslinking RNA to nylon filters

This protocol is to be used only when a Stratalinker or other calibrated crosslinking device is not available.

1. Wash filters post-Northern transfer as described above and blot excess liquid by placing the RNA-containing filter membrane on top of one or two sheets of Whatman 3MM paper. Do not allow filters to dry out completely; it is best to irradiate filters that are still damp.
2. Place filter(s) face down directly on the surface of the transilluminator, the surface of which has been wiped clean just prior to use.

Caution: Be sure to wear proper eye protection, because serious permanent damage can occur to the eyes and skin if not properly shielded.

3. The appropriate exposure time is a direct function of the wavelength and age of the transilluminator. With 302 nm-emitting instruments, a total of 45 s is usually more than adequate. For 254 nm-emitting instruments, crosslinking is complete in about 30 s.

Note 1: These are general guidelines only. The exact parameters must be empirically determined for each instrument.

Note 2: In this laboratory, filters are irradiated for one-half of the estimated time required, rotated 90°, and then irradiated again for the remaining interval (don't forget to turn off the transilluminator before attempting to move the filter). Rotation is suggested because of the parallel orientation of the bulbs within the transilluminator, some of which may be burned out in older instruments. Depending on the exact specifications of a particular instrument, this may or may not be necessary.

4. Store filters sandwiched between Whatman 3MM paper in a dry location out of direct light until further use.

For most applications, UV crosslinking, which generally requires less than 1 min to perform, has superseded baking in a vacuum oven for 2 h. In addition to the crosslinking application, calibrated UV light sources can be used to nick stained RNA or DNA gels to facilitate complete blotting. Often, the amount of UV light to which the samples are exposed during photodocumentation is sufficient in this regard.

Post-fixation handling of filters

The time frame for Northern analysis changes dramatically after transfer and immobilization have been accomplished. It should be abundantly clear by now that up to this point all manipulations involving labile RNA should revolve around the actual date of RNA isolation. This pertains to poly(A)⁺ purification (if necessary), determination of concentration and purity, electrophoresis, blotting, nuclease protection analysis (Chapter 15), conversion into complementary DNA (cDNA; Chapter 17), and so forth. Once a sample is immobilized on a

solid support (filter membrane), it can be stored for months until hybridization (or rehybridization) is convenient.

The important thing to remember is that filters that have yet to be subjected to hybridization should be stored dry or in a vacuum oven to prevent moisture and opportunist growth from encroaching. Filters that have been subjected to hybridization, which will be rescreened, should be washed at very high stringency to remove the old probe as soon after completing hybridization detection as possible. Do not allow filters to dry out until the previously hybridized probe has been removed. Until then, keep the filter damp and wrapped in plastic. It is extremely difficult to completely remove hybridized probe from filters that have been allowed to dry.

Reverse northern analysis

The traditional format of Northern analysis is the measurement of the abundance, albeit with limited sensitivity, of one transcript at a time. The rise of assays designed to monitor the abundance of dozens, hundreds, or thousands of genes simultaneously, primarily by PCR-based approaches, was the driving force behind the development of newer methods to *confirm* purported changes in gene expression. Data from this laboratory has been met quite favorably by journal reviewers when purported changes in gene expression from quantitative and semi-quantitative PCR experiments are confirmed using a non-PCR-based approach; this can be a most worthwhile publication strategy! The reverse Northern analysis, or simply the “reverse Northern”, is one such approach (Zhao *et al.*, 1996; Zhang *et al.*, 1997; Kang *et al.*, 1998). In one variant of this technique, DNA (usually cDNA) sequences representing genes that are believed to be expressed differentially are spotted directly onto duplicate nylon filters in a dot-blot-like format⁵. Then, each filter is hybridized to a pool of radiolabeled total cellular RNA (the probe) from either the control or experimental samples. In a more common variant of this technique, the duplicate filters are hybridized to a pool of total cDNA (the probe) from the control and experimental samples, respectively. Usually the cDNA probes are synthesized as needed by performing first-strand cDNA synthesis on the original RNA samples; this latter approach precludes potential RNase-associated difficulties when working with an RNA probe. Less cDNA in one of the probes is a consequence of having less mRNA in the original biological sample, and will produce proportionately lower signal intensity upon hybridization (Fig. 11.3). Confirmation of differential gene expression occurs when the hybridization signal from a specific cDNA is stronger on one filter compared to the other filter. For even greater certainty, some investigators elect to blot each cDNA sequence in two (usually adjacent) locations on the filter to ensure the authenticity of hybridization signal strength.

⁵Described in Appendix N.

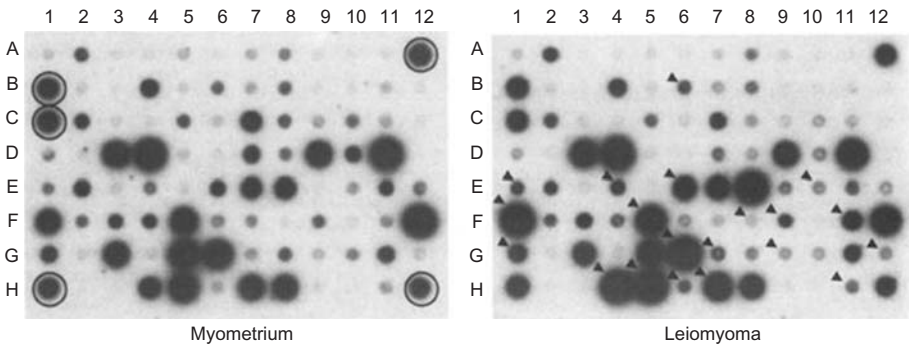


Figure 11.3 Reverse Northern analysis. Two sets of blots are made with an equal quantity of PCR products. Positive controls (G3PDH, also known as GAPDH) are marked with circles. Notice the asymmetry with respect to placement of the controls. Genes that have been up-regulated are indicated with arrows. Leiomyoma is a benign smooth muscle tumor and myometrium is normal smooth muscle tissue from the wall of the uterus. Reprinted by permission from Macmillan Publishers Ltd: *Cell Res.* 12, 215–221, copyright 2002.

The advantage and convenience of the reverse Northern procedure is that one may choose to analyze only a few sequences at a time. This method is also much, much less expensive than the microarray approach and is ideal for laboratories that are interested only in genes associate with a specific biological phenomenon, such as apoptosis or cancer. Considering the way that microarray analysis is performed (Chapter 21), one might think of that type of global analysis of gene expression as a glorified form of reverse Northern analysis!

References

- Alwine, J.C., D.J. Kemp, and G.R. Stark. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci.* 74, 5350–5354.
- Alwine, J.C., D.J. Kemp, B.A. Parker, J. Reiser, J. Renart, G.R. Stark, and G.M. Wahl. (1979). Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. *Methods Enzymol.* 68, 220–242.
- Arnhem, N. and E.M. Southern. (1977). Heterogeneity of ribosomal genes in mice and men. *Cell* 11, 363–370.
- Bittner, M., P. Kupferer, and C.F. Morris. (1980). Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzyloxymethyl cellulose or nitrocellulose sheets. *Analyt. Biochem.* 102, 459–471.
- Chomczynski, P. (1992). One-hour downward alkaline capillary transfer for blotting DNA and RNA. *Analyt. Biochem.* 201, 134–139.
- Chomczynski, P. and P.K. Qasba. (1984). Alkaline transfer of DNA to plastic membrane. *Biochem. Biophys. Res. Commun.* 122, 340–344.

- Church, G.M. and W. Gilbert. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci.* **81**, 1991–1995.
- Kang, D.C., R. LaFrance, Z.Z. Su, and P.B. Fisher. (1998). Reciprocal subtraction differential RNA display: An efficient and rapid procedure for isolating differentially expressed gene sequences. *Proc. Natl. Acad. Sci.* **95**, 13788–13793.
- Khandjian, E.W. (1987). Optimized hybridization of DNA blotted and fixed to nitrocellulose and nylon membranes. *Biotechnology* **5**, 165–167.
- Li, J.K., B. Parker and T. Kowalik. (1987). Rapid alkaline blot-transfer of viral dsRNAs. *Analyt. Biochem.* **163**, 210–218.
- Löw, R. and T. Rausch. (1994). Sensitive, nonradioactive Northern blots using alkaline transfer of total RNA and PCR-amplified biotinylated probes. *BioTechniques* **17**(6), 1026–1029.
- Medveczky, P., C.W. Chang, C. Oste, and C. Mulder. (1987). Rapid vacuum drive transfer of DNA and RNA from gels to solid supports. *BioTechniques* **5**(3), 242–246.
- Olszewska, E. and K. Jones. (1988). Vacuum blotting enhances nucleic acid transfers. *Trends Genet.* **4**, 92–94.
- Pitas, J.W. (1989). A simple technique for repair of nylon blotting membranes. *BioTechniques* **7**(10), 1084.
- Reed, K.C. and D.A. Mann. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13**, 7207–7221.
- Saito, I., H. Sugiyama, N. Furukawa and T. Matsuura. (1981). Photochemical ring opening of thymidine and thymine in the presence of primary amines. *Tetrahedron Lett.* **22**, 3265–3268.
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.
- Zhang, H., R. Zhang, and P. Liang. (1997). Differential screening of differential display cDNA products by reverse Northern. *Methods Mol. Biol.* **85**, 87–93.
- Zhao, S., S.L. Ooi, F.C. Yang, and A.B. Pardee. (1996). Three methods for identification of true positive cloned cDNA fragments in differential display. *BioTechniques* **20**, 400–404.

12 Nucleic Acid Probe Technology

Rationale	261	Polymerase chain reaction	274
Probe classification	263	Random priming	274
Selection of labeling system	263	Nick translation	274
Isotope labeling	265	5' End-labeling	275
Minimizing decomposition problems	266	3' End-labeling	275
Non-isotopic labeling	267	Antisense RNA probes	276
The ubiquitous dyes Cy3 and Cy5	267	characteristics of RNA probes	278
Popular chemiluminescence formats	268	RNA probe synthesis	278
Biotin	268	<i>In vitro</i> transcription	278
Digoxigenin	270	5' End-labeling	279
Fluorescein	271	3' End-labeling	279
Direct enzyme labeling	271	Probe purification	279
DNA probes	272	Probe storage	280
DNA probe synthesis	273	References	281

Rationale

At the heart of molecular biology is the base-pairing of two complementary nucleic acids molecules to form a double-stranded structure. This process is known as hybridization. In each hybridization event, each of the strands has a different name. In classical molecular biology techniques, such as Northern analysis and Southern analysis, one partner of the hybridization pair is known as the **target**; this strand is traditionally immobilized on a filter membrane, as described in the previous chapter. The other partner in the hybridization pair, known as the **probe**, carries some type of label that allows it to be followed throughout the experiment. The probe usually circulates in a warmed solution of defined pH and ionic strength (hybridization solution) and will base-pair with any complementary sequences that have been immobilized on the surface of the filter. This is known as mixed-phase hybridization, in which the experiment culminates with the localization and quantification of the probe on the filter by autoradiography or by chemiluminescence.

Most probe molecules are significantly shorter than the target molecules to which they hybridize. In fact, probe molecules can be as short as 20 bases and still maintain an ability to discriminate between closely related sequences from organisms with complex genomes. The fact that hybridization probes can see so

short affords the investigator great latitude in probe design, especially given the prevalence of- and easy access to genome and transcriptome data at GenBank and other such sites. Shorter probes tend to hybridize to complementary target sequences faster than longer probes. One enormous advantage to using short probes is that the hybridization can be performed in aqueous (formamide-free) hybridization buffers. Finally, short probes can be custom-designed based on available amino acid sequence information, particularly if no probe is available. In this approach, very small amounts of protein sequence information are necessary: a 20-mer corresponds to about 7 amino acid codons. Probe design using this strategy is described in detail in Chapter 18.

Hybridization can also occur in a format in which both strands are free floating: base-pairing between probe and target is driven by random collision kinetics. Since neither strand is fixed in place on any filter, hybridization is driven to completion more rapidly than is observed in the mixed-phase format. The absence of the filter increases the sensitivity of these assays, usually by a factor of at least ten. Examples of solution hybridization techniques include the nuclease protection (Chapter 15) assays as well as PCR (Chapter 18). The reason for the increase in sensitivity is due to the fact that the act of fixing nucleic acid molecules on a filter renders some of them incapable of hybridization, perhaps due to their orientation with respect to the filter or else due to the mechanics of filter fixation.

The selection of the type of nucleic acid probe best suited for a particular application is just as important as the methodology by which hybridization events will be localized and quantified. The role of a probe, ideally, is to hybridize to every complementary sequence present in a hybridization reaction. Nucleic acid probes have numerous applications, owing in no small measure to their powerful discriminatory abilities. They can be used to detect, among other things (1) quantitative or qualitative changes in gene expression; (2) gene amplifications and deletions; (3) gene rearrangements; (4) chromosomal translocations; (5) point mutations; and (6) the presence of new genetic sequences in cells (pathogens). Moreover, oligonucleotides, universally used as primers to support the polymerase chain reaction, were originally used as nucleic acid probes for blot analysis.

As stated previously, a nucleic acid probe is a long or short polynucleotide that carries some type of label or tag. The label allows the investigator to follow the probe throughout the experiment and provides a means for at least semi-quantitative detection. Historically, probe synthesis (i.e., the incorporation of label) and detection were entirely dependent on the efficient incorporation of radiolabel accompanied by sensitive autoradiographic techniques. The major decision to be made was the method by which dNTP-radiolabeled precursors were to be incorporated. Now, however, the more pressing question in each new study is: "When can non-isotopic labeling and detection be used in place of radiolabeling and autoradiography?"

The not-so-direct answer is "Often, but much depends on the precise application and required level of sensitivity." For the most part, non-isotopic procedures are quite compatible with nearly all molecular biology applications involving sample fixation on filter membranes (not nitrocellulose). These techniques

include, but are not limited to, Northern analysis, Southern analysis, Western analysis, dot and slot blot analyses, colony hybridization, plaque screening, and non-filter-based *in situ* hybridization. The rule of thumb that this Author frequently proffers is: "If you can see bands by autoradiography after the third day under film¹, then it is likely that you can switch to chemiluminescence without loss of sensitivity or resolution. Autoradiographic analysis requiring more than three days may or may not be fully compatible with non-isotopic methods."

Some of the other non-filter-based techniques that offer enhanced sensitivity, including nuclease protection assays and the nuclear runoff assay, are best performed with radiolabel. In many cases PCR-based innovations have superseded the need to perform some of these older techniques at all.

Probe classification

Nucleic acid probes can be either homogeneous or heterogeneous in nature. Homogeneous and heterogeneous probes may be either DNA or RNA molecules. In a hybridization solution, a probe preparation is said to be homogeneous if all probe molecules are the same. For example, one would use only *c-fos* oncogene cDNA molecules to quantify the abundance of *c-fos* transcripts in a sample. Heterogeneous probes consist of mixtures of two or more sequences that may be closely related in nucleotide sequence or may be completely dissimilar. For example, in the preparation of a subtraction library, the goal is to identify mRNA sequences that are unique to one population of cells by subtracting or removing all sequences expressed in common with a control population of cells. Thus, the probe might be made of thousands of different types of members. Another example is an instance where several oligonucleotides are present in a hybridization reaction because of sequence ambiguities pertaining to the target molecules of interest. When planning to work with heterogeneous probes, however, one should avoid targets that are known to contain moderately- and highly repetitive sequences (e.g., *Alu*) so as not to generate gross and unsatisfactory hybridization signals.

Selection of labeling system

The key considerations in the selection of a probe labeling system are (1) required sensitivity and resolution; (2) whether the probe is DNA or RNA; (3) method of label incorporation, (4) probe stability after labeling; (5) type of hybridization; (6) desired method of detection.

There are three basic methods for labeling nucleic acids: (1) radiolabeling; (2) hapten incorporation (biotin, fluorescein, digoxigenin) that will support

¹Autoradiography performed at -80° with an intensifying screen.

non-isotopic detection by chemiluminescence, chemifluorescence, or chromogenic techniques; and (3) direct enzyme labeling. These techniques are compatible with the labeling of DNA, RNA, and (usually) oligonucleotides for most applications. The methods, merits, and drawbacks of the permutations are described here. Further, numerous systems are available commercially for synthesis of nucleic acid probes, and each has a recommended detection methodology for optimal sensitivity and resolution.

Probes are classified depending on the distribution of label with the probe molecule itself. End-labeled probes are those to which the label has been added at the 5' end, as with the kinasing reaction (Fig. 12.1a), or at the 3' end by the action of terminal deoxynucleotidyl transferase or simply "terminal transferase" (Fig. 12.1b). Continuously-labeled probes are those in which the label has been added along the length of the backbone of the molecule itself, at fairly regular intervals (Fig. 12.2). Examples of continuously labeled probes include, but are not limited to, those generated by random priming, the polymerase chain reaction (PCR), *in vitro* transcription, or cross-linkage of a hapten or enzyme directly to the backbone of a molecule. The choice between end-labeling and continuously-labeling techniques depends on (1) whether the probe is DNA or RNA; (2) whether the probe is single-stranded or double-stranded; (3) whether the probe, if double-stranded, has a recessed or protruding 5' and 3'

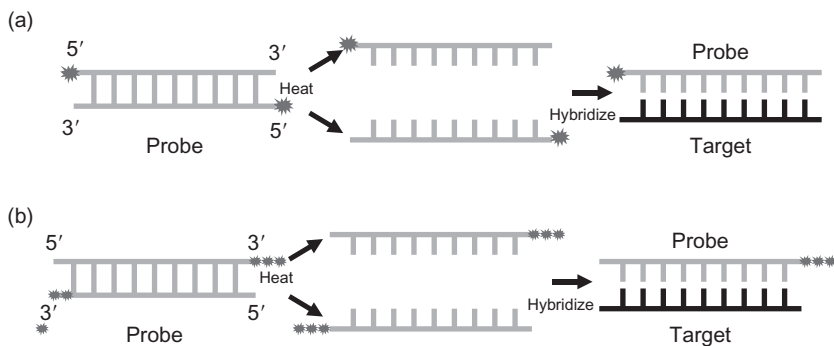


Figure 12.1 DNA end-labeling procedures for probe synthesis. (a) The kinasing reaction transfers a radiolabel to the 5' phosphate. (b) Terminal transferase is used to label the 3' end of DNA probe molecules.

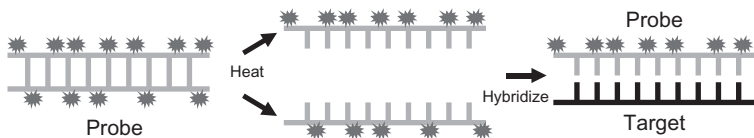


Figure 12.2 Continuous probe labeling. Incorporation of label along the length of the DNA probe can be accomplished by any of several methods, including random priming, nick translation, and PCR.

end(s); (4) the required level of label incorporation and detection sensitivity; (5) the length of the probe to be labeled; and (6) whether the probe, if double-stranded DNA, is linear or a covalently closed circle (a plasmid). Each of these permutations is addressed in detail below.

Isotope labeling

CAUTION: Care and common sense should always be used when working with or storing radioactive materials. A lab coat, disposable gloves, eye protection, film badges to monitor exposure, and proper training are essential at all times, regardless of the isotope or its specific activity. Be certain to adhere closely to departmental or institutional regulations concerning shielding, handling, disposal, and laboratory safety when working with or around or any other hazardous material. In every case, the radiation safety officer should be consulted for specific applications and institution-wide policies.

In many applications, isotopes offer excellent sensitivity and compatibility with many labeling techniques. Some of the inherent disadvantages of isotope labeling include the half-life of the isotope, potential health hazards, containment of radioactivity, purchase cost, disposal cost, probe stability and usefulness after labeling, and the need for relatively long detection periods. For the convenience of the reader, the half-life data for commonly used isotopes are presented in Table 12.1.

The most frequently used isotopes for nucleic acid probe synthesis are ^{32}P , ^3H , and ^{35}S and, less frequently, ^{33}P . Nucleotide triphosphate precursors are labeled in the appropriate position (α or γ), thereby supporting enzyme-mediated transfer of the portion of the nucleotide containing the isotope to the molecules being labeled. The intended application and the precise method of detection dictate the selection of isotope. For example, the long path length associated with β -emission from ^{32}P -labeled probes makes this label useful for recording hybridization events on X-ray film, and for precisely the same reason, anatomical resolution is inadequate for RNA target localization *in situ*. This necessitates the use of an alternative labeling and detection system.

The ready availability of ^{33}P has assuaged some of the dangers inherent in and fears associated with the use of ^{32}P . ^{33}P -labelled nucleotides are an alternative

Table 12.1 Isotope Decay Data

Isotope	β Emax	Physical half-life	Amount remaining after 30 days	Amount remaining after 60 days	Amount remaining after 1 year
^{32}P	1.71 MeV	14.29 days	0.233	0.054	0.0
^{33}P	0.249 MeV	25.34 days	0.441	0.195	0.00005
^{35}S	0.167 MeV	87.4 days	0.788	0.621	0.05541
^3H	0.019 MeV	12.28 years	0.995	0.991	0.94534

for use in general molecular biology applications, especially DNA sequencing (Zagursky *et al.*, 1991). The half-life of ^{33}P is 25.4 days with a β E_{max} of 0.25 MeV, which is about seven-fold less than that of ^{32}P ($E_{\text{max}} = 1.71$ MeV). ^{33}P can be handled on the bench-top using routine radiation safety practices but with less bulky shielding (Zagursky *et al.*, 1991) than is required when handling ^{32}P , though one should always exercise caution nonetheless and follow guidelines set for by the Radiation Safety Office at your location.

Minimizing decomposition problems

The shelf-life of a nucleotide depends on two factors: its radioactive half-life and its biological half-life. Whereas the radioactive half-life is a constant, the rate of biological degradation varies depending on the handling of the material and the temperature. Storing vials of radioactive nucleotides in a refrigerator or worse, at room temperature, can greatly increase the rate of radiolysis and biological degradation. Decomposition products consist primarily of nucleotide monophosphates and inorganic phosphate, although nucleotide diphosphates can also be present. Nucleotides are shipped to arrive on dry ice and should be stored frozen below -20° , or at -80° if possible. Storage in frost-free freezers is not recommended.

The following is a partial list of recommendations from to help radiochemical users minimize decomposition problems. More complete details regarding the safe handling and usage of radiochemicals can found in PerkinElmer Life Sciences' excellent "Nuclide Safe Handling" link which can be accessed on-line at www.perkinelmer.com/lifesciences, and from which several of the following points have been abstracted².

1. Use the radiochemical as soon as possible after receiving it. Prolonged storage allows time for additional nuclear decay, causing increased radical generation, and therefore, greater amounts of decomposition. Such decomposition will occur even if the radiochemical container is left unopened and stored under the most ideal conditions.
2. Store the sample properly. Pay close attention to the recommended storage conditions, as outlined on the technical data sheet that accompanies each product.
3. Follow the technical data sheet instructions when opening an ampoule containing radiochemicals. Refer to these instructions for storage information.
4. Thaw, at room temperature, until melting, ^3H -labeled radiochemicals, stored in aqueous solvents, and ^{35}S -, ^{125}I - and ^{32}P -labeled biochemicals, except ^{35}S - and ^{32}P -labeled nucleotides, which should be quick-thawed. For radiochemicals stored at very low temperatures, slowly thawing the sample in the refrigerator (or on ice) is recommended.
5. Do not allow radiochemicals to sit at room temperature for prolonged periods of time, because this often increases the rate of radiochemical decomposition significantly.
6. Minimize the number of times the primary container is opened, and reseal the container immediately after each use. If a compound will be used several times, aliquot

²Reproduced with permission.

the required amounts into separate vials for storage. Each time a solution is handled, impurities, especially oxygen and water, may be introduced.

7. Use clean syringes and pipette tips when withdrawing aliquots from the primary container, since many buffers and salts used in biological systems are harmful to the radiochemical.

Non-isotopic labeling

An ever-increasing variety of non-isotopic labeling and detection methodologies have been developed in an attempt to minimize the use of radioisotopes and the inherent dangers associated with their presence in the laboratory. Truly amazing refinements have been made in many of these systems, facilitating extraordinary levels of detection without any radioactivity at all. In many cases, the sensitivity achievable is equivalent to that of ^{32}P , though much depends on the patience of the user. Comparatively speaking, the various non-isotopic labeling and detection systems offer similar levels of sensitivity and resolution, though they differ markedly in the number of different steps which must be performed by the user, especially pertaining to the detection component of the system. Among the more common non-isotopic labeling methods are biotinylation, labeling with digoxigenin (DIG), fluorescein labeling, direct enzyme labeling, or labeling with fluorescent dyes. Biotin, DIG, and fluorescein labeling support detection by chromogenic (colorimetric) methods and by chemiluminescence. Because of the intrinsic technical difficulties in many filter-based chromogenic techniques, this type of detection chemistry never gained widespread popularity. Now, however, efficient non-isotopic labeling and detection technologies have become standard fare in molecular biology laboratories. Although the notion of eliminating radioactive probes is enticing, to say the least, the investigator must also be aware that the mechanics of non-isotopic labeling, probe purification, and detection may be drastically modified, compared to methods associated with the use of "hot" probes. The most significant change that experienced ^{32}P users will notice is the amount of dedicated time that must be allocated to perform the detection procedure. Usually, the posthybridization stringency washes are followed by a blocking step. Also, once the chemiluminescence substrates are applied, the filter cannot be washed at higher stringency unless one repeats the series of posthybridization washes and blocking steps in its entirety. The benefits reaped by the elimination of isotopes, from a safety and economics perspective, far outweigh the additional attention that this type of detection requires.

The ubiquitous dyes Cy3 and Cy5

Although the reagents needed to perform the classical methods of non-isotopic labeling and detection are still available from select vendors, a new generation of fluorescent labels, and strategies for incorporating these labels, have been developed and are in widespread use. Among the current favorites are the

water-soluble fluorescent dyes Cy3 (indocarbocyanine; emits at 570 nm) and Cy5 (indodicarbocyanine; emits at 670 nm). These dyes are used for labeling a variety of biomolecules, including nucleic acids. These dyes can be used to label DNA probes for nucleic acid hybridization by direct label incorporation via labeled nucleotides, or by indirect labeling via aminoallyl dUTP or aminoallyl dCTP, followed by a secondary reaction to link Cy3 or Cy5 to the probe. The most common contemporary application for the use of these and other fluorescent nucleotides is the synthesis of cDNA for microarray analysis (Chapter 21).

Popular chemiluminescence formats

There are several ways to perform labeling reactions that support detection by chemiluminescence. Some of these methods involve hapten-labeling the probe, while another popular format is known as a direct enzyme labeling. In the case of the former method, an enzyme that supports chemiluminescence is introduced posthybridization, during the detection protocol, and in the case of the latter, the enzyme is already present throughout the hybridization and stringency washes, after which the appropriate substrates are introduced immediately. Each approach, as describe below, has advantages and disadvantages.

Biotin

Biotin is a small water-soluble vitamin that can be readily incorporated into a number of biological molecules. For probe synthesis, biotinylation can be accomplished in any one of several ways:

1. Biotinylated nucleotides. Virtually all of the classical methods for enzymatic probe synthesis, including random priming, nick translation, and PCR, can be performed using biotinylated nucleotides, specifically Bio-11-dUTP³, which functions as a dTTP analog. The prefix Bio means that the nucleotide is biotinylated and that there is an 11-atom spacer arm separating the biotin itself from the nucleoside component. Bio-21-dUTP is also available for labeling DNA, as are Bio-11-UTP⁴ and Bio-16-UTP for labeling RNA. In general, the longer the linker arm, the greater the ease with which the biotin can interact with the required streptavidin during the detection process, described below. In contrast, the shorter the linker arm, the more efficient is the incorporation of the labeled nucleotide by nick translation, random priming, or PCR. Bio-11-dUTP seems to offer the best balance between efficiency of label incorporation and detection sensitivity and is thus widely used for non-isotopic DNA probe synthesis.
2. Biotinylated primers. The easiest method of biotinylation is during the actual synthesis of oligonucleotides, which can be used directly as probes or to support the polymerase chain reaction. Primers can be biotinylated at their 5' end during synthesis at minimal additional cost. One need not worry about biotinylation later

³Biotin-11-2'-deoxyuridine-5'-triphosphate, tetralithium salt.

⁴Biotin-11-2'-uridine-5'-triphosphate.

- on, and the biotin label will not interfere with primer specificity or amplification efficiency.
3. Photoactivatable biotin acetate⁵ (Forster *et al.*, 1985; McInnes *et al.*, 1987; Denman and Miller, 1989). This material, an aryl azide derivative of biotin, can be used to label single- or double-stranded DNA or RNA. The labeling reaction is driven by visible light and generally requires 20 to 30 minutes to complete. Biotin incorporation occurs, on average, only once in every 50 to 100 nucleotides, thereby minimizing its usefulness for labeling oligonucleotides. Probes labeled with photoactivatable biotin acetate are as stable as any other biotin-labeled nucleic acid molecules. It should be noted, however, that photoactivatable biotin has fallen out of favor since the development of photo-reactive TFPA-PEG₃-Biotin (Pierce Chemical Company), which offers improved sensitivity and versatility.
 4. Psoralen biotin (Cimino *et al.*, 1985; Wassarman, 1993). Psoralens and their derivatives are planar tricyclic compounds that intercalate into double-stranded nucleic acids upon irradiation with 320 to 400 nm light. These molecules become covalently attached to nucleotide bases, preferentially to thymidines, resulting in a much higher degree of label incorporation, compared to labeling with photoactivatable biotin. To a lesser extent, single-stranded molecules, such as oligonucleotides, may also become modified. Psoralen biotin derivatives are very useful for the synthesis of biotinylated probes for nucleic acid hybridization. Unincorporated psoralen biotin conjugates are removed by simple extraction with *n*-butanol.

After labeling, biotinylated probes are stable for at least 1 year when stored at -20° . An aliquot of probe is thawed just prior to use and, if double-stranded, is boiled, rapidly cooled on ice, and then added immediately to the hybridization buffer. Biotinylated probes support both chromogenic detection and chemiluminescence.

From a detection point of view, it really does not matter how the biotin becomes part of the probe, though 5' biotinylation of oligonucleotides is one of the easiest and most efficient methods to accomplish labeling. Biotinylation almost never interferes with biological activity and, in the case of biotinylated nucleotides such as Bio-11-dUTP⁶ (Fig. 12.3), linker arms between the biotin itself and the backbone of the probe effectively minimize steric interference and, from a stability perspective, the fewer the freeze-thaw cycles, the better.

Biotinylated probes are localized posthybridization via the strong natural affinity observed between biotin and streptavidin ($K_d = 10^{-15}M^{-1}$). Streptavidin is a tetrameric protein (MW 60,000), isolated from the bacterium *Streptomyces avidinii* and has four biotin binding sites. Unlike avidin (from egg), streptavidin has neutral isoelectric point at physiological pH, with few charged groups,

⁵Photobiotin acetate: *N*-(5-azido-2-nitrophenyl)-*N'*-(3-biotinylaminopropyl)-*N'*-methyl-1,3-propanediamine. Available from Pierce Biotechnology, Rockford, IL.

⁶dUTP is widely used as a non-isotopically labeled nucleotide, functioning as a dTTP analog. In study after study, the use of dUTP consistently generated the highest levels of label incorporation, by comparison with any of the other four deoxynucleotides in biotinylated form, using the standard molecular biology DNA polymerases. Moreover, both AMV and MMLV reverse transcriptase will likewise recognize these modified nucleotides for the synthesis of biotin-labeled, first-strand cDNA. RNA probes can also be synthesized via *in vitro* transcription in the presence of Bio-11-UTP, DIG-11-UTP, or Fl-12-UTP.

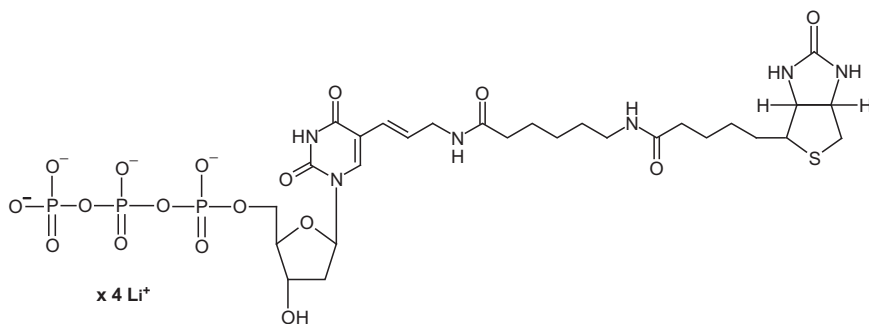


Figure 12.3 Biotin-11-dUTP.

and contains no carbohydrate. These properties reduce non-specific binding and background problems, thereby enhancing the sensitivity of many forms of this assay. Many biotin-streptavidin applications in nucleic acid hybridization detection and related techniques have been described (Leary *et al.*, 1983; Wilchek and Bayer, 1984; Hofman and Finn, 1985). In some systems, the streptavidin used to bind biotin is a modified conjugate, usually to alkaline phosphatase (AP), though some systems feature a streptavidin-horseradish peroxidase (HRP) conjugate. This positions the enzyme needed to support the actual detection process itself. Depending on the specific substrate to which the AP or HRP is subsequently exposed, hybridization events are then localized and quantified by light emission captured on X-ray film (chemiluminescence) or the formation of a color precipitate directly on the filter (chromogenic detection).

Digoxigenin

Digoxigenin (DIG; Roche Applied Science) is a steroid hapten derived from plants of the genus *Digitalis* (*D. lanata*, and the purple foxglove *D. purpurea*), the latter of which is same source as the cardiac medication digitalis. It is widely used as nucleic acid label and supports both chromogenic detection and chemiluminescence (Martin *et al.*, 1990). DNA probes are enzymatically labeled, usually by random priming with digoxigenin-dUTP (DIG-11-dUTP; Fig. 12.4), and RNA probes are synthesized by *in vitro* transcription with DIG-11-UTP. These nucleotides are linked via a spacer arm to DIG. The resulting DIG-labeled molecules then function as hybridization probes in much the same manner as any other type of probe. Oligonucleotides can also be DIG-labeled during their synthesis, thereby precluding cumbersome, inefficient labeling reactions at some later date. As with biotinylation, DIG labeling does not interfere with biological activity, and probes so labeled are stable at -20° for at least 1 year after labeling. DIG labeling supports posthybridization detection by chemiluminescence or by formation of an insoluble color precipitate directly on the filter membrane. This method is also well respected for *in situ*

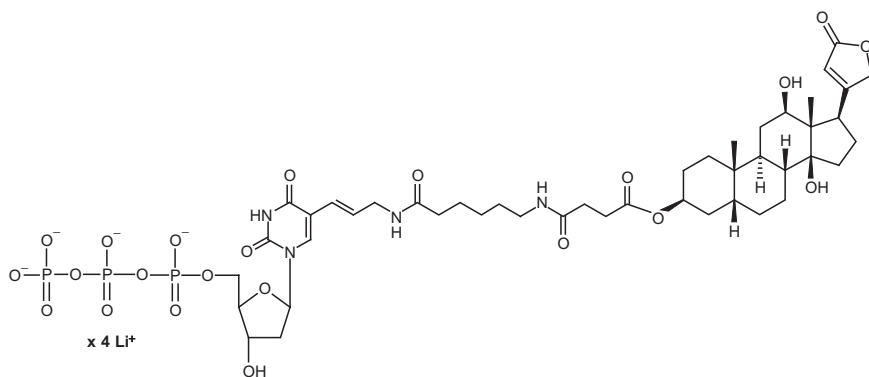


Figure 12.4 DIG-11-dUTP.

hybridization applications (Robbins *et al.*, 1991; Crabb *et al.*, 1992; Schröder *et al.*, 2000).

Following the posthybridization stringency washes, DIG-labeled probes are detected by enzyme-linked immunoassay, using an antibody conjugate (anti-DIG-alkaline phosphatase). As with systems involving biotin, the emission of light or the formation of precipitate is mediated by alkaline phosphatase (AP)-mediated dephosphorylation of a substrate compatible with the method of detection, either colorimetric or chemiluminescence. Horseradish peroxidase (HRP)-labeled antibody conjugates are also in use.

Fluorescein

As recently as a few years ago, the only fluorescent nucleotides readily available for either DNA or RNA probe synthesis were fluorescein (FL)-labeled: FL-12-dUTP for the synthesis of DNA probes and FL-12-UTP for the synthesis of RNA probes (Fig. 12.5). These nucleotides were used to generate probes in much the same way as biotin- and DIG-labeled nucleotides or by direct incorporation during oligonucleotide (primer) synthesis. Following hybridization, high-affinity antibodies prepared against fluorescein were used to localize the probe, and the conjugation of these antibodies with either AP or HRP could support either chemiluminescence or chromogenic detection. The intrinsic fluorescence of the hapten could also be exploited to monitor the incorporation of label into the probe, thereby allowing the investigator to assess the efficiency of the labeling reaction.

Direct enzyme labeling

Each of the previously described techniques for generating DNA probes involves the substitution or addition of labeled nucleotides to the probe. In contrast to these time-honored methods, direct enzyme labeling (ECL System; GE Healthcare) involves crosslinking an enzyme, capable of supporting

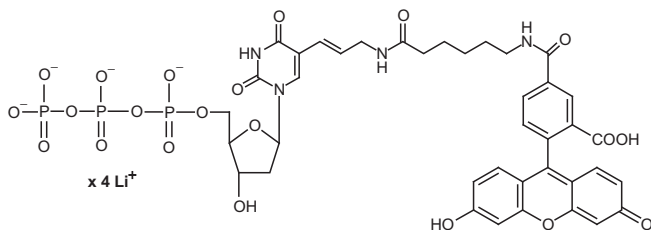


Figure 12.5 FL-12-dUTP.

chemiluminescence, directly to the backbone of heat-denatured DNA, using glutaraldehyde as a bridge or linker arm. The method is extremely efficient and requires only a total of 20 min to label the probe. Further, there are no subsequent steps required for the removal of unincorporated label. Succinctly, double-stranded DNA is boiled for 5 min and then rapidly cooled on ice to minimize renaturation. Then, glutaraldehyde links a modified HRP to the now single-stranded DNA. Because the probe cannot be boiled again after the labeling has been completed, and because the denatured DNA will, of course, reanneal in a relatively short time, one should label only the required mass of probe, which should be added to the hybridization mix as soon as possible after labeling. Because the enzyme is present throughout the hybridization and subsequent washes (next chapter), stringency must be regulated by modifying ionic strength and by the inclusion of destabilizing agents such as SDS and urea, rather than by raising the temperature. In this laboratory the resulting X-ray films almost always of outstanding quality with very little background.

DNA Probes

DNA probes can be a diverse lot. The following examples of DNA probes include double-stranded cDNA or genomic DNA sequences, first-strand cDNA synthesized directly from a mixture of mRNAs in the presence of labeled dNTPs, oligonucleotides, and PCR products:

1. Double-stranded probes require denaturation into their constituent single strands to be used as probes, and a plethora of clones and sequence information is widely available. A description of a number of labeling techniques follows. One disadvantage of DNA probes is that they are less thermodynamically stable than RNA probes (see Chapter 8 for commentary on stringency), though this often has negligible impact on the outcome of an experiment.
2. First-strand cDNA can be synthesized from mRNA (or from an enriched fraction thereof) in the presence of labeled dNTP precursors. This will generate a heterogeneous probe that can be used directly for hybridization. Heterogeneous probes are particularly useful for the identification of sequences that are uniquely present in one of two or more RNA populations (subtraction hybridization, Chapter 22), for microarray screening, and for other applications.

3. Oligonucleotides are artificially synthesized single-stranded DNA. "Oligos" were used as standard probes for hybridization for years before PCR was developed. Thus, the primers that an investigator might now be using to support amplification of a sequence of interest by PCR could easily function as stand-alone probes for standard nucleic acid hybridization. Oligonucleotides are classified as "long" oligomers, consisting of as many as 100+ bases, or "short" oligomers, usually consisting of fewer than 30 bases. The advantage of working with oligonucleotides is that the investigator has complete control over the sequence, which is made to order. The main disadvantage of using oligonucleotides is the same as the main advantage: because the investigator has control over the sequence of the oligonucleotide, one must know *which* sequence to have synthesized. The short oligonucleotide probes offer the greatest flexibility and experimental latitude in terms of discriminating power and predictable thermodynamic behavior.

In the event that a definitive oligonucleotide probe sequence is not discernible, perhaps due to codon degeneracy, it is possible to use several oligonucleotides simultaneously in what might be thought of as a hybridization cocktail; alternatively, one may elect to use only one oligonucleotide, though at a lower stringency, by lowering the hybridization and washing temperatures. While this strategy reduces the concentration of the correct complementary sequence and may favor semi-non-specific hybridization, the important point is that authentic hybridization events do not go undetected. See Chapter 18 for a description of how degenerate oligonucleotide probes and primers should be designed.

4. Any PCR product can be used as a probe. Label incorporation can occur during the PCR process by incorporation of hapten-labeled nucleotides into the PCR products. Alternatively, primers can be synthesized with a hapten at the 5' end; because the primers become part of the product, label incorporation is efficient and automatic during amplification of the template.

DNA probe synthesis

A variety of methodologies for labeling DNA have been described. In short, these methods are used to generate end-labeled or continuously labeled probes. Most enzyme-mediated labeling techniques are very much dependent on a polymerase activity, which is responsible for incorporation of the labeled nucleotides. Further, the use of *Taq* or other thermostable DNA polymerases permits labeling reactions to be performed at higher temperatures via PCR, thereby reducing the incidence of enzyme-mediated point mutations during probe synthesis. Of course, one must also consider the natural error rate of the *Taq*. Proofreading enzymes and enzyme blends, described in Chapter 18, may be helpful in this regard.

A brief description of the more common techniques follows. All of these labeling reactions are commercially available in kit form, containing all necessary reagents, with the exception of isotopes for radiolabeling techniques. In general, most labeling systems used to generate radioactive probes can also be used to generate a variety of non-isotopic probes. When radiolabeling, be sure that the position of the isotope in the nucleotide precursor (i.e., α or γ label) supports the intended labeling method.

Polymerase chain reaction

PCR is an excellent method for probe synthesis, requiring exquisitely small quantities of template material. In the presence of the appropriate radiolabeled or hapten-labeled nucleotide precursor, PCR products are labeled as they are being synthesized. Alternatively, the primers themselves may be labeled non-isotopically during their own synthesis, negating the requirement for the inclusion of labeled nucleotide precursors as part of the reaction mix.

PCR as a labeling method offers several advantages: speed, versatility, efficiency of the reaction, and the fact that most labs are now performing PCR routinely. Minute quantities of starting material are required, and the resulting molecules are of uniform length. When performed in the presence of labeled nucleotides, PCR-generated probes are continuously labeled and generally show a high degree of label incorporation. For comprehensive reviews of reaction parameters and other PCR-related strategies, see Dieffenbach and Dveksler (2003); Innis *et al.* (1999); McPherson and Hames (1995); Sambrook and Russell (2001).

Random priming

Random priming is a type of primer extension in which a mixture of small oligonucleotide sequences, acting as primers, anneal to a heat-denatured double-stranded template (Feinberg and Vogelstein, 1983; 1984). The annealed primers ultimately become part of the probe itself, because the Klenow fragment of DNA polymerase I extends the primers in the 3' direction and, in so doing, incorporates the label. Random priming works significantly better with linearized DNA molecules; attempts to label covalently closed, supercoiled DNA typically result in probe specific activities 20 to 30-fold less than those of the corresponding linear DNA. DNA molecules between 200 and 2000 bp are the best candidates for random priming, although template length is not a critical parameter in this labeling reaction. When labeling with ^{32}P , random priming typically produces probes with specific activities of about $1\text{--}3 \times 10^9$ cpm/ μg when 2.5–10 ng of starting template is used. Labeling usually requires 10–30 min.

Nick translation

Nick translation (Rigby *et al.*, 1977) is one of the oldest probe labeling techniques. It involves randomly nicking the backbone of a double-stranded DNA with dilute concentrations of DNase I. At extremely low concentrations, this enzyme nicks a template at four or five sites, producing a free 3'-OH group that can act as a primer at each nicking location. Next, the enzyme DNA polymerase I removes the native nucleotides from the probe molecules in the 5'→3' direction (exonuclease activity) while replacing them with labeled dNTP precursors by virtue of its 5'→3' polymerase activity. Nick translation is efficient for both linear and covalently closed DNA molecules, and labeling requires about 1 h. When labeling with ^{32}P , nick translation produces probes with specific activities of approximately $3\text{--}5 \times 10^8$ cpm/ μg .

5' End-labeling

One alternative method to generating continuously labeled probes is to label the 5' end of the molecule with the addition of a radiolabeled phosphate. This method of 5' end-labeling is colloquially known as the kinasing reaction; it specifically involves the transfer of the γ phosphate of ATP (not dATP) to a 5'-OH substrate group on dephosphorylated DNA molecules (forward reaction). This labeling reaction is catalyzed by the enzyme T4 polynucleotide kinase and is an excellent method for labeling short oligonucleotides. One common error in the laboratory pertaining to this labeling reaction involves the purchase of α -labeled nucleotides (needed for continuously labeled and 3'-end-labeled probes) rather than the required γ -labeled ATP.

The forward kinasing reaction is far more efficient than the exchange reaction, which involves the substitution of 5' phosphates, so it is prudent to remove the naturally occurring 5' phosphate group first, using a phosphatase. DNA kinasing reactions typically require 30 min, and this type of labeling is renowned for generating extremely high specific activity probes. Finally, a variety of non-isotopic labeling kits are available from several vendors for 5'-end modification.

3' End-labeling

Probe synthesis by 3' end-labeling involves the addition of nucleotides to the 3' end of either DNA or RNA. DNA 3' end-labeling is most often catalyzed by terminal transferase. Single- and double-stranded DNA molecules are labeled by the addition of dNTP to 3'-OH termini; double-stranded, blunt-ended fragments and fragments with 3' overhang structure are labeled most efficiently. A most unusual feature of terminal transferase is that there is no template strand requirement, i.e., this enzyme is a non-template-dependent polymerase⁷. Consequently, any one dNTP (or dNTP mixture) can be selected for addition to the 3' end of a polynucleotide. This type of labeling produces a 3' overhang of five or six extra nucleotides. For some applications, these non-blunt ends could then be polished with the Klenow fragment of DNA polymerase I or T4 DNA polymerase to support blunt end ligation although, depending on the precise end-structure of the molecules, the loss of label may occur.

The disadvantage of this labeling technique can be pronounced when the labeling of short oligonucleotides is involved. The addition of nucleotides changes the length of the molecule, which may or may not modify the specificity of the oligonucleotide, as well as its thermodynamic characteristics. If an oligonucleotide is labeled by nucleotide addition to the 3' end, it is strongly suggested that dATP be used, because of the lesser thermodynamic stability of A::T base-pairing compared to G::C base pairs.

⁷The *Taq* DNA polymerase also has terminal transferase activity, as do other non-proofreading polymerases: it is a well known fact that *Taq* polymerase adds an extra nucleotide, usually an "A", to the 3' end of each polynucleotide strand that it synthesizes. Double-stranded PCR products made by *Taq*, therefore, are expected to have an extra "A" at the 3' end of each strand.

Antisense RNA Probes

RNA probes as hybridization tools remain popular because of several key advantages associated with their use. These probes are synthesized by *in vitro* transcription and can be substituted for DNA probes in nearly all applications. In the construction of the transcription template, the cDNA to be transcribed is usually ligated between two different transcription promoters from bacteriophage, which flank a multiple cloning site and are positioned in opposite orientations. The most common constructions feature the highly efficient SP6, T7, and/or T3 RNA polymerase promoters (Fig. 12.6). This construction is linearized with an appropriate restriction enzyme prior to initiating the transcription reaction, favoring the efficient synthesis of large amounts of uniform length, continuously labeled probe (Fig. 12.7).

An integral part of the preparation of RNA probes is the transcription of antisense and sense transcripts from the same template construction. Antisense RNA is complementary to mRNA, and is therefore able to base pair to it. In addition to the ability to function as a probe in transcription studies, antisense RNA can inhibit gene expression at the translational level by forming a non-translatable double-stranded RNA structure: this is the very heart of RNA interference (RNAi). Antisense RNA was long ago shown to regulate gene expression *in vitro* in mammalian systems (Nishikura and Murray, 1987), in phages, bacteria, and plants (Green *et al.*, 1986), and in animals (Knecht and Loomis, 1987). RNAi is discussed in detail in Chapter 23. Antisense probes are widely used in the area of *in situ* hybridization.

Example

An investigator wishes to characterize the histopathology of a diseased rat cerebellum by identifying which cells, if any, are synthesizing *N-myc* mRNA.

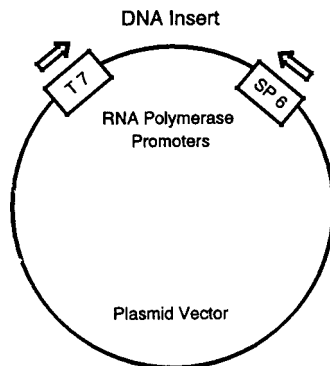


Figure 12.6 Typical orientation of dual RNA polymerase promoters in a plasmid construction that supports *in vitro* transcription. The template for transcription, usually cDNA, is cloned between the two promoters. This plasmid construction is then linearized prior to initiation of the *in vitro* transcription reaction.

To make this determination, a tissue section is prepared for *in situ* hybridization⁸. Because the interest of this study is mRNA, antisense RNA (cRNA) should be used as the probe. The enhanced thermodynamic stability of cRNA when hybridized to mRNA permits very stringent assay conditions. Sense RNA probe, that is, probe with the same sequence as the mRNA, is also generated and used as a negative control to assess the degree of non-specific hybridization. To support chromogenic detection, sense and antisense probe could be synthesized in the presence of DIG-labeled UTP (Roche Applied Science), followed by chromogenic detection. The distribution of color precipitate is assessed by light microscopy and requires no isotope whatsoever. After the distribution of message *in situ* is discerned, antisense RNA can be used to obtain more quantitative information pertaining to the expression of *N-myc* using PCR based methods.

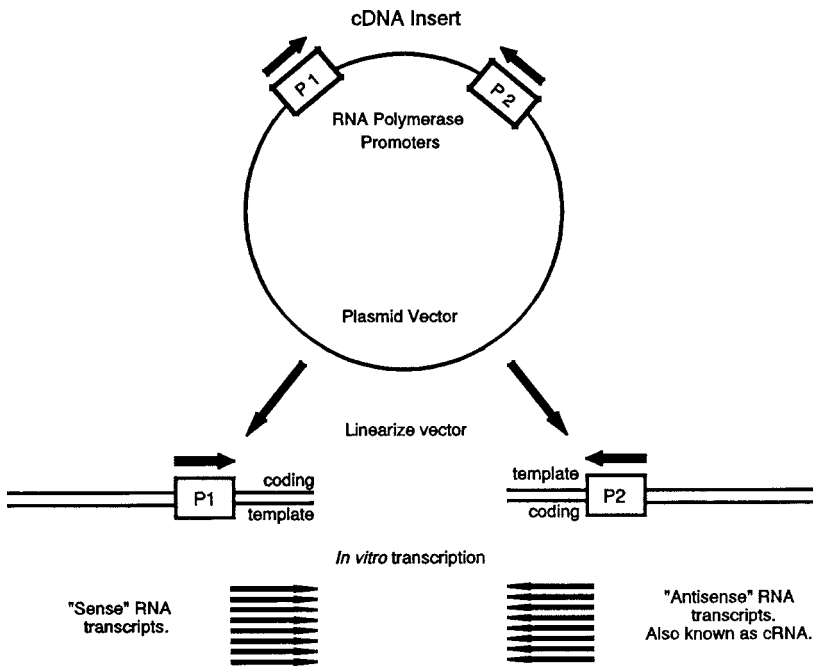


Figure 12.7 *In vitro* transcription templates must be linearized prior to the transcription reaction, thereby favoring the efficient production of large quantities of continuously labeled probe of uniform length. Note that the positioning of the promoters influences whether the resulting transcript will be sense or antisense RNA. If the orientation of the insert cDNA is unknown, then empirical determination will be required to discern which promoter gives which type of transcript.

⁸For excellent discussions of *in situ* hybridization, see Chesselet (1990), Wilkinson (1998), Darby and Hewitson (2006), Nuovo (1994), Polak and McGee (1998).

Characteristics of RNA probes

1. RNA probes are single-stranded. While such probes do not require boiling prior to use, heating them briefly will help to perturb any intramolecular base-pairing that may have occurred.
2. All RNA probe molecules are available for hybridization. Single-stranded, denatured RNA molecules cannot renature, as do dsDNA, although some intramolecular base-pairing, should it occur, may reduce the effective concentration of the probe.
3. RNA probes are continuously labeled as they are being transcribed, thereby generating probes with a very high degree of label incorporation.
4. RNA probes show greater thermodynamic stability when base-paired with either DNA or RNA target molecules, compared to the thermodynamic stability associated with DNA probes.
5. RNA probes are synthesized by *in vitro* transcription from a linearized template; therefore, all probe molecules are of uniform length.
6. Enormous quantities of probe can be synthesized in a single *in vitro* transcription reaction⁹.
7. The SP6, T7, and T3 bacteriophage RNA polymerase promoters demonstrate virtually no cross-reactivity; therefore, transcription reactions initiated from one promoter or the other are virtually free of transcripts of the opposite sense. Consequently, both sense and anti-sense RNA probes can be synthesized as needed.
8. RNA transcribed *in vitro*, as with all RNA, must be treated with RNase-free reagents; failure to do so will result in rapid degradation of the probe.
9. RNA probes often produce unacceptable, high levels of background; thus, at the conclusion of the hybridization period it is a common practice to digest all probe molecules with RNase A and RNase T1. This treatment will result in degradation of all probe molecules that did not participate in duplex formation.
10. Radiolabeled RNA probes are often of such high specific activity that they may experience radiolysis if stored for extended periods.

RNA probe synthesis

Compared to the diverse methods for DNA probe synthesis, there is only one reliable method for labeling RNA probes, namely *in vitro* transcription. Because of the intrinsically labile nature of RNA and the susceptibility to RNase degradation, RNA probes must be treated with the same care as any other RNA preparations. A brief description of RNA labeling techniques follows. *In vitro* transcription systems are available in kit form from many vendors and contain all necessary reagents (except isotopes). These systems work well in the non-isotopic format as well.

In vitro transcription

In vitro transcription is the only reliable, economical method for generating RNA probes. Large amounts of efficiently labeled probes of uniform length can

⁹The proportion of full-length transcripts can be increased by lowering the standard reaction temperature (Krieg and Melton, 1987).

be generated by transcription of a DNA sequence ligated next to an RNA promoter. One excellent strategy is to clone the DNA to be transcribed between two promoters in opposite orientations. This allows either strand of the cloned DNA sequence to be transcribed in order to generate sense and antisense RNA for hybridization studies. This can be accomplished via the use of universally available cloning vectors. Alternatively, a template sequence can be positioned between one or two transcription promoters via PCR. RNA probes synthesized by *in vitro* transcription are referred to generically as Riboprobes[®], even though this is a trademark protected name (Promega).

RNA probes synthesized by transcription *in vitro* are assembled from NTP precursors, much as occurs *in vivo*. Transcripts are elongated by the addition of nucleotide monophosphates into the nascent backbone of the probe. Isotopic labeling with ³²P-UTP requires radiolabel in the α position, as with all continuously labeled probes. Since the probe is labeled as it is synthesized, the degree of label incorporation is very high.

5' End-labeling

One alternative method to generating continuously labeled RNA probes by *in vitro* transcription is to label the 5' end of the molecule with the addition of a radiolabeled phosphate. This method of 5' end-labeling is colloquially known as the kinasing reaction; it specifically involves the transfer of the γ phosphate of ATP to a 5'-OH substrate of RNA or DNA (forward reaction). The forward kinasing reaction is far more efficient than the exchange reaction which involves the substitution of 5' phosphates. This labeling reaction is catalyzed by the enzyme T4 polynucleotide kinase. The relative efficiency of this method is many-fold greater for DNA than for RNA, however, which is one reason why it is not commonly utilized to label RNA.

One common error in the laboratory pertaining to this labeling reaction involves the purchase of α -labeled nucleotides (needed for continuously labeled and 3' end-labeled probes) rather than the required γ -labeled ATP.

3' End-labeling

RNA can also be 3' end-labeled using the enzyme poly(A) polymerase. This enzyme, which is naturally responsible for nuclear polyadenylation of many hnRNAs, catalyzes the incorporation of AMP. Isotopic labeling requires α -labeled ATP precursors. In addition to its utility in RNA probe synthesis reactions, poly(A) polymerase can be used to polyadenylate naturally poly(A)⁻ mRNA and other RNAs in order to support oligo(dT) primer-mediated synthesis of cDNA.

Probe purification

As with most labeling techniques, unincorporated precursor must be separated from the labeled probe. Failure to do so, especially when working with

isotopes, usually results in unacceptably high levels of background. There are three basic methods for cleaning up the probe. For optimal results, follow the manufacturer's recommendations that accompany the labeling system.

In the first approach, the probe can be ethanol precipitated, in which case the unincorporated label remains in the supernatant. Whereas this is an efficient approach for purifying larger probes, attempted precipitation of oligonucleotides can be rather difficult, especially if small masses of nucleic acids are involved. Precipitation of probes is accomplished with a combination of salt and alcohol (see Chapter 2, Table 2.2), most often 0.1 volume 3 M NaOAc and 2.5 volumes of 95% ethanol.

The second general method for probe purification is gel filtration, or a permutation thereof. One popular approach is to make or purchase a so-called spun column (Sambrook and Russell, 2001), which consists of little more than a 1 ml syringe packed with Sephadex (GE Healthcare); Sephadex G-50 is used for larger probes, while Sephadex G-25 is reserved for the purification of labeled oligonucleotides. Spun-column chromatography is a glorified form of gel filtration, involving centrifugation of the column at 1000 to 1200 \times g for approximately 4 min; the probe is rapidly eluted while the unincorporated nucleotides lag behind in the column. In this lab, more than 95% of the unincorporated label is routinely removed after the labeling reaction. Gel filtration probe purification is compatible with a wide variety of labeling methods, both isotopic and non-isotopic.

The third general method for probe purification involves the use of one of the numerous concentration or absorbent devices for probe purification. These devices include the High Pure PCR system (Roche Applied Science), Elutip-d[®] minicolumns (Whatman), Ultrafree[®] filtration units, and Centricon[®] devices (Millipore, Billerica, MA), to name but a few. All represent rapid, convenient alternatives to traditional column chromatography or alcohol precipitation, for separating radiolabeled nucleic acids from unincorporated nucleotides. The disposable nature of these devices minimizes the spread of radiolabeled waste products as well.

Probe storage

Isotopically labeled probes are useful only as long as there is sufficient activity remaining and the probe molecules themselves remain intact. In general, radiolabeled probes should be used as soon after labeling as possible. Radiolysis of probes, especially those labeled to extremely high specific activity, becomes more problematic as the probe ages. Store probes at -20° or -80° until just prior to use. Avoid repeated freezing and thawing of probes.

Non-isotopic labeling systems vary with respect to the recommended post-labeling storage and handling of probes. Biotinylated, DIG-labeled, and fluorescein-labeled probes can be stored at -20° for up to 1 year. In some labs, an entire day is invested generating non-isotopic probes for the experiments that are planned

for the next several weeks or months. Probes are then available for hybridization reactions as needed. Probes should never be denatured until just before use.

References

- Chesselet, M.F. (1990). *In Situ Hybridization Histochemistry*. CRC Press, Boca Raton, FL.
- Cimino, G.D., H.B. Gamper, S.T. Isaacs and J.E. Hearst. (1985). Psoralens as photoreactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Ann. Rev. Biochem.* **54**, 1151–1193.
- Crabb, I.D., S.S. Hughes, D.G. Hicks, J.E. Puzas, G.J. Tsao and R.N. Rosier. (1992). Nonradioactive in situ hybridization using digoxigenin-labeled oligonucleotides. Applications to musculoskeletal tissues. *Am. J. Pathol.* **141**, 579–589.
- Darby, I.A. and T.D. Hewitson. (2006). *In Situ Hybridization Protocols: Isotopic and Non-isotopic Detection Methods*. Humana Press.
- Dieffenbach, C. and G.S. Dveksler (Eds). (2003). *PCR Primer*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Denman, R.B. and D.L. Miller. (1989). Use of photobiotinylated deoxyoligonucleotides to detect cloned DNA. *BioTechniques* **7**, 138–141.
- Feinberg, A.P. and B. Vogelstein. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **132**, 6–12.
- Feinberg, A.P. and B. Vogelstein. (1984). Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **137**, 266–267.
- Forster, A.C., J.L. McInnes, D.C. Skingle and R.H. Symons. (1985). Non-radioactive hybridization probes prepared by chemical labeling of DNA and RNA with a novel reagent, photobiotin. *Nucleic Acids. Res.* **13**, 745–761.
- Green, P., O. Pines and M. Inouye. (1986). The role of antisense RNA in gene regulation. *Ann. Rev. Biochem.* **55**, 569–597.
- Hofman, K. and F.M. Finn. (1985). Receptor affinity chromatography based on the avidin–biotin interaction. *Annals N.Y. Acad. Sci.* **447**, 359–372.
- Innis, M.A., D.H. Gelfand and J.J. Sninsky. (1999). *PCR Applications: Protocols for Functional Genomics*. Academic Press, San Diego, CA.
- Knecht, D.A. and W.F. Loomis. (1987). Antisense RNA inactivation of myosin heavy chain gene expression in *Dictyostelium discoideum*. *Science* **236**, 1081–1086.
- Krieg, P.A. and D.A. Melton. (1987). *In vitro* synthesis with SP6 RNA polymerase. *Methods Enzymol.* **155**, 397–415.
- Leary, J.J., D.J. Brigati and D.C. Ward. (1983). Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose. *Proc. Natl. Acad. Sci.* **80**, 4045–4049.
- Martin, R., C. Hoover, S. Grimme, C. Grogan, J. Holtke and C. Kessler. (1990). A highly sensitive, nonradioactive DNA labeling and detection system. *BioTechniques* **9**, 762–768.
- McInnes, J.L., S. Dalton, P.D. Vize and A.J. Robins. (1987). Non-radioactive photobiotin labeled probe detects single copy genes and low abundance mRNA. *Nature BioTechnology* **5**, 269–272.
- McPherson, M.J. and B.D. Hames (Eds). (1995). *PCR 2: A Practical Approach*. IRL Press at Oxford University Press, Oxford.

- Nishikura, K. and J.M. Murray. (1987). Antisense RNA of proto-oncogene *c-fos* blocks renewed growth of quiescent 3T3 cells. *Mol. Cell. Biol.* 7, 639–649.
- Nuovo, G.J. (1994). *PCR in situ Hybridization*. Raven Press, New York, NY.
- Polak, J.M. and J.O. McGee. (1998). *In Situ Hybridization: Principles and Practice*. Oxford University Press, Oxford.
- Rigby, P.W.J., M. Dieckmann, C. Rhodes and P. Paul. (1977). Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113, 237–251.
- Robbins, E., F. Baldino, J.M. Roberts-Lewis, S.L. Meyer, D. Grega and M.E. Lewis. (1991). Quantitative non-radioactive *in situ* hybridization of preproenkephalin mRNA with digoxigenin-labeled cRNA probes. *Anat. Rec.* 231, 559–562.
- Sambrook, J. and D.W. Russell. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schröder, S. M. Hain and K. Sterflinger. (2000). Colorimetric *in situ* hybridization (CISH) with digoxigenin-labeled oligonucleotide probes in autofluorescent hyphomycetes. *Internat. Microbiol.* 3, 183–186.
- Wassarman, D.A. (1993). Psoralen crosslinking of small RNAs *in vitro*. *Mol. Biol. Rep.* 17, 143–151.
- Wilchek, M. and E.A. Bayer. (1984). The avidin–biotin complex in immunology. *Immunol. Today* 5, 39–43.
- Wilkinson, D.G. (Ed.). (1998). *In Situ Hybridization: A Practical Approach*, 2nd edition. Oxford University Press, Oxford.
- Zagursky, R.J., P.S. Conway and M.A. Kashdan. (1991). Use of ³³P for Sanger DNA sequencing. *BioTechniques* 11, 36–38.

13 Practical Nucleic Acid Hybridization

Rationale	283	T_m for long probes	291
Factors influencing hybridization kinetics and duplex stability	284	T_m for oligonucleotide probes	292
Temperature	285	Hybridization and the Northern analysis	292
Ionic strength	286	Prehybridization: filter preparation	293
pH	287	Protocol: prehybridization for long probes	293
Probe length	287	Probe denaturation	294
Probe concentration	287	Hybridization	295
G+C content	288	Posthybridization stringency washes	295
Mismatching	288	Probe removal and rehybridization	296
Probe complexity	289	Protocol: generic method for probe removal	297
Viscosity	289	References	298
Formamide	290		
Urea	290		
Hybridization temperature	291		

Rationale

Perhaps the most complex, yet least understood component of molecular biology studies are the parameters that govern nucleic acid hybridization, a process first described by Marmur and Doty (1961). At the heart of molecular hybridization, two complementary polynucleotide molecules hybridize in an antiparallel fashion to form a double-stranded molecule. This phenomenon is also known as base-pairing, duplex formation, annealing, and renaturation.

These days, most investigators are familiar with the concept of hybridization in the context of the polymerase chain reaction, and all of its variants, in which complementary nucleic acid molecules are repeatedly coaxed to hybridize and then denature for the purpose of amplification. Thus, a profound understanding of the thermodynamic behavior of the specific molecules involved (the primers and the template) is necessary to design a temperature cycling profile that will support high fidelity amplification. Moreover, one must understand that hybridization kinetics are influenced by other components of the reaction, and these variables are discussed here.

In the context of traditional blot analysis, and under conditions that promote nucleic acid hybridization, the strands that participate in duplex formation are known by specific names. The **probe** strand, as the name implies, usually carries some type of label that permits localization and quantitation of the probe at the conclusion of the hybridization period. The other strand, commonly known as the **target**, either may be immobilized on a solid support for filter-based analysis (e.g., Northern analysis) or may simply be suspended in buffer, free to participate in solution hybridization (e.g., S1 nuclease analysis).

For example: the objective of a study might be the assessment of *c-myc* transcription in cells acquired by biopsy. Because the sample contains thousands of different mRNA species, addressing this question would require a suitable complementary probe, that is, one that would be able to hybridize only to *c-myc* transcripts. The probe itself could be a cDNA sequence, a genomic sequence, an antisense RNA sequence, or an oligonucleotide. The *c-myc*-specific probe might be labeled with ^{32}P , for detection by autoradiography; alternatively the probe could be hapten-labeled or prepared by direct enzyme labeling, using any of the techniques described in Chapter 12 that support detection by chemiluminescence. Assuming that the hybridization stringency is satisfactory, *c-myc* probe molecules will only hybridize to complementary *c-myc* transcripts, which are the target molecules for that particular probe. When performed properly, the probe is always present in a large molar excess so as to ensure that all *c-myc* transcripts will be hybridized with probe molecules.

At the completion of the hybridization, the filter is characteristically washed to remove unbound, left-over probe. The magnitude of hybridization is then assessed in a manner consistent with the labeling method used to generate the probe at the onset of the experiment, either by autoradiography, by chemiluminescence, or by fluorescence. Further, to permanently record the location of size standards on the X-ray film, the investigator may wish to label a very small amount of probe that will be able to hybridize to the size markers, making detection of these nucleic acid species possible as well. Having the size markers visible on the film eliminates the guesswork regarding the relative distribution of size standards compared to experimental samples.

The expression of other genes can also be studied, requiring only a different probe for each gene of interest. In fact, one strategy is the sequential hybridization of a filter containing experimental RNA: after the first round of hybridization and detection, the probe is removed by very high stringency washes and then the filter is hybridized with a different probe that will assay for the presence of a different type of mRNA. Nylon membranes are very well-suited for this approach.

Factors influencing hybridization kinetics and duplex stability

The original Southern analysis protocol (Southern, 1975) described the hybridization of a DNA probe to DNA target sequences. The resulting DNA:DNA

duplex molecules are, in fact, thermodynamically less stable than the DNA:RNA or RNA:RNA duplexes that form when RNA targets, RNA probes, or both, are involved. The point is that the stringency at which the hybridization and posthybridization washes are performed must be attuned to the chemical nature of the duplexes that form between probe and target molecules: DNA:DNA versus DNA:RNA versus RNA:RNA.

Example

An investigator may wish to screen a zebra library using a llama cDNA probe. It may well be that a related sequence does not exist in the zebra genome. Alternatively, a similar sequence might be present but because of sequence divergence, the llama probe might not be able to hybridize stably under stringent hybridization conditions. To identify zebra sequences in a library that are related to the llama cDNA probe, or to further characterize such sequences by Northern analysis or Southern analysis, it would be necessary to lower the stringency, as discussed in Chapter 8, in order to hybridize based on partial complementarity.

Factors that influence the rate, specificity, fidelity, and probable utility of hybridization probes include, but are not limited to: temperature, ionic strength (primarily Na^+), pH, the organic solvent formamide, urea, guanine and cytosine (G+C) content of probe, probe length, probe concentration, probe complexity (the total length of different probe sequences present during the hybridization), degree of complementarity between probe and target sequences, the degree of mismatching, and the viscosity of the system. It is important to realize that the influence of each of these variables is also dependent on the mobility of the probe and target molecules, specifically (1) solution hybridization, where both probe and target molecules move randomly in an application-specific hybridization cocktail, as in the nuclease protection assays and PCR; or (2) mixed-phase hybridization in which the target sequences are immobilized on a solid support, as in Northern analysis.

Temperature

Perhaps the most frequently and easily manipulated variable that can either promote or prevent hybridization is the temperature of the system. A measure of the stability of hybrids and the extent to which they are expected to form can be predicted by calculating their melting temperature (T_m). In the context of molecular hybridization, the T_m is best thought of as an equilibrium point. The T_m is that temperature at which 50% of all possible hybrids can form, while 50% remain dissociated into their constituent single strands: for each new duplex that forms at the T_m , another duplex is expected to dissociate. Historically, this phenomenon has been characterized by monitoring UV light absorption at 260 nm (A_{260}) as the temperature of a DNA solution

is gradually raised. As a double-stranded DNA begins to melt (denature) into its constituent single strands, there is an accompanying increase in UV light absorption, known as a hyperchromic shift. This type of assay is useful as a means of characterizing the thermodynamic stability of double-stranded molecules in general. This is not to be confused with melting curve analysis that is commonly performed in association with real-time PCR (Chapter 19). The hyperchromic shift phenomenon described here is an intrinsic property of nucleic acid molecules and is measured without the need for any fluorescent DNA binding dyes such as SYBR Green or other fluorescent dyes. By measuring the hyperchromic shift associated with double-stranded DNA, the T_m of the DNA species is determined as that temperature where A_{260} is exactly half way between the absorbance maximum and absorbance minimum; at this (equilibrium) temperature 50% of the DNA is denatured and 50% remains double-stranded.

The maximum rate of hybridization for long probes is typically observed 15 to 20° below the calculated T_m in a system. In aqueous salt solutions, this corresponds to a hybridization temperature of about 60 to 65°. The inclusion of formamide modifies this parameter as described below. With respect to characterizing nucleic acids from one species using a probe derived from another species, a 1° change in melting temperature corresponds to about 1% sequence divergence (Bonner et al., 1973). The inclusion of formamide, as is quite common, permits significant reduction of hybridization temperature without a relaxation of stringency.

Last, since only 50% of all possible hybrids form at the T_m of the probe, and it is unlikely that the investigator wants only 50% of all target molecules hybridized, the actual hybridization temperature (T_{HYB}) is usually at least 5° below the T_m , and perhaps quite a bit more. This is especially true if there is reason to believe that the probe is not exactly matched with the target. The idea is to get the probe to base-pair; the reality of the situation is that the investigator has far more control over the stringency of the assay in the washing steps, conducted at the conclusion of the hybridization period and before initiating the detection procedures. So powerful are the posthybridization washing steps that in some cases hybridization can be performed at room temperature; as long as the filters are washed correctly prior to probe detection, a relatively low temperature hybridization would be unlikely to have a negative impact on the outcome of the experiment.

Ionic strength

Stringency is dramatically influenced by increasing or decreasing the amount of salt in a hybridization buffer. In general, the rate of hybridization increases with salt concentration up to about 1.2M NaCl, at which point the rate becomes constant; the T_m of a nucleic acid duplex changes approximately 16° with each factor of 10 in salt concentration (Britten and Davidson, 1985). Further, because divalent cations have a greater influence than monovalent cations

at low concentrations, restrained use of chelating agents such as EGTA and EDTA is strongly recommended¹ Typical hybridization buffers usually range between 0.1–1.0 M Na⁺.

pH

One of the factors that profoundly influences the stability of double-stranded molecules is the pH of the environment. Most hybridization reactions are conducted at near neutral pH because alkaline buffers promote duplex dissociation, and highly acidic buffers may well result in depurination of both probe and target molecules. In general, pH is among the least frequently manipulated variables in nucleic acid hybridization.

Probe length

The length of the probe is directly factored into the calculation of the T_m of the probe (duplex stability), according to the expression $D = 500/L$ where D is the reduction in T_m (in degrees Celsius) and L is the number of base pairs actually participating in duplex formation (Britten and Davidson, 1985), which is usually the length of the probe itself. The shorter the probe, the more rapidly hybridization occurs and the more discriminating the probe becomes. For example, short oligonucleotides can discriminate between closely related target sequences that differ by as little as one base. As a general rule, the shorter the probe, the more influential many of the other variables become.

Probe concentration

The investigator is afforded considerable leeway with respect to the amount of probe added to a hybridization solution. In short, as the concentration of probe increases, the forward hybridization kinetics also increases. Thus, higher probe concentrations accelerate hybridization, reducing the amount of time needed to saturate all of the target sequences. This is especially true when oligonucleotides are used as probes. Be aware, however, that the use of excessive quantities of probe can also favor non-specific hybridization. Think about what happens in PCR, for example, when excessive primer concentrations are used – the outcome can be disastrous.

To identify the proper probe concentration, two variables must be considered: (a) the method of detection, that is, isotopic or non-isotopic; and (b) the probable abundance of the target, for example, single-copy genes or low abundance transcripts. The amount of probe required should be based on the guidelines shown in Table 13.1 and is based either on the mass of the probe or, in the case of radiolabeling, on the specific activity of the probe. As it is important

¹EGTA = ethylene-bis(oxyethylenenitrilo)tetraacetic acid. EDTA = ethylenediaminetetraacetic acid.

Table 13.1 Recommended Nucleic Acid Probe Concentrations

	Probe mass (ng/ml)	Probe activity (cpm/ml)
Isotopic labels		
Single-copy genes or low abundance transcripts	20–50	10 ⁶
Multiple-copy genes or high abundance transcripts	5–20	10 ⁵
Oligonucleotide probes	1–5	>10 ⁴
Non-isotopic labels		
Single-copy genes or low abundance transcripts	50–75	Not applicable
Multiple-copy genes or high abundance transcripts	20–50	Not applicable
Oligonucleotide probes	5–10	Not applicable

Note: Values indicated are per milliliter of hybridization solution. Probe concentration can be expressed either by actual mass added or by label incorporation. While detection sensitivity with non-isotopic probes is often comparable to that with radiolabeling, the amount of probe added per milliliter of hybridization solution is usually greater. One may compensate for lower probe concentrations by simply extending the hybridization period.

that the probe be present in a large molar excess, the stringency of the assay is readily controlled during the posthybridization washing steps.

G+C content

The stability of a hybrid molecule is profoundly influenced by the base composition of the molecules involved. Because three hydrogen bonds occur between guanine (G) and cytosine (C), GC-rich duplexes are thermodynamically more stable than adenine/thymine (AT)-rich duplexes, which have only two hydrogen bonds between them. The importance of G+C content becomes more pronounced as the probe involved becomes shorter. For example, in the design of oligonucleotide probes, both for hybridization and for PCR, the paramount importance of the G+C content in duplex stability is underscored by the expression:

$$T_m = 0.41(\%GC) + 69.3$$

where (%GC) is the percentage of the duplex consisting of guanine and cytosine residues (Marmur and Doty, 1962). This relationship holds in 1× SSC buffer (150 mM NaCl; 15 mM Na₃-citrate, pH 7.0).

Mismatching

Many of the recommended hybridization conditions that accompany the purchase of filter membranes and molecular biology kits assume a perfect or nearly perfect match between probe and target sequences. Probes that manifest

this quality are known as exact-match probes which, under stringent conditions, rapidly form duplexes with the target. A mismatch is a position in a probe where the base is unable to form hydrogen bonds with the base in the corresponding location in the target molecule. The implication is that most or all of the other bases along the length of a probe are able to base-pair to target. A mismatch could be the result of a point mutation in the organism under investigation, a naturally occurring polymorphism, or an error in probe design. Mismatching retards the hybridization rate and at sufficiently high temperatures, mismatches may inhibit hybridization altogether. With longer probes (greater than 1 kb) the effect of mismatch may not be noticeable under standard hybridization conditions². This mismatch effect is significantly more pronounced with shorter probes, especially when using oligonucleotides as probes or as primers to support PCR. To promote hybridization between mismatched probes and target sequences, the stringency of the system can be relaxed, usually by decreasing the hybridization temperature. The more the hybridization temperature is reduced below the T_m of the probe, the more mismatch between probe and target will be tolerated. Below a certain sequence-dependent temperature, however, random hybridization will occur.

Probe complexity

With respect to hybridization, the complexity of the probe is the length of *different* probe sequences involved in the hybridization that could potentially base-pair to complementary target molecules. In most cases, only a single probe sequence would be used (e.g., a *c-myc* probe to assess the prevalence of cytoplasmic *c-myc* mRNA by Northern analysis). When using degenerate oligonucleotide probes, however, in which more than one probe sequence is present (perhaps due to target sequence ambiguity), the complexity is greater, the concentration of the correct oligonucleotide probe is reduced, and hybridization kinetics are slowed. When the complexity of the probe is very large, meaning that there is very little sequence redundancy, it might be useful to divide the various probe sequences into pools, based on T_m similarities, and perform multiple hybridizations and stringency washes at temperatures that are compatible with the probes in a particular pool or batch. This approach is described in greater detail in Chapter 18.

Viscosity

The original studies of the effect of viscosity on nucleic acid hybridization (Thrower and Peacocke, 1968; Subirana and Doty, 1966; Chang et al., 1974)

²Standard hybridization conditions may consist of any of a variety of multipurpose prehybridization and hybridization buffers and often feature an overnight incubation. "Standard" implies that the effect of probe length is not a major variable, and abatement of the rate of hybridization due to mismatches is probably masked by the relatively long period of hybridization. As the length of the probe decreases, hybridization conditions become less "typical."

showed a reduction in the rate of renaturation as viscosity (presence of sucrose, glycol) increases. However, anionic polymer dextran sulfate, a common ingredient in many older hybridization buffer formulations (Wahl et al., 1979, Lederman, 1981), also contributes to the viscosity of the milieu but appears to increase the rate of hybridization because the dextran sulfate occupies physical space in the hybridization cocktail, thereby forcing the probe into a smaller volume. This increases the effective concentration of the probe. A similar phenomenon occurs with the inclusion of Denhardt's solution³ (Denhardt, 1966) in various hybridization recipes. It is an excellent alternative to dextran sulfate and alleviates the severe background problems frequently associated with the inclusion of dextran sulfate in many hybridization recipes.

Formamide

Formamide (HCONH₂) destabilizes double-stranded molecules by interfering with hydrogen bond formation. Thus, the inclusion of freshly deionized formamide in hybridization recipes allows a reduction in T_m (and hybridization temperature) in a linear manner by about 0.75–1.0° for each 1% of added formamide. This roughly corresponds to a T_m reduction of roughly 2.4–2.9° per mole of formamide, depending on the G+C content and other variables (Blake and Delcourt, 1996). It is therefore commonplace to perform nucleic acid hybridization in buffers consisting of 50% formamide at 42°; this corresponds approximately to $T_m - 20^\circ$ for a “typical” probe. Thus, maximum hybridization between complementary nucleic acid molecules is favored with no relaxation of stringency. Conducting RNA analysis in formamide-containing buffers also extends the useful life of probe molecules that may be heat labile. One should be aware, however, that formamide is rather volatile at the temperatures commonly used to perform hybridization, so care should be taken to avoid exposure to its toxic fumes.

The reader should be aware that a new generation of non-formamide-based hybridization buffers, which drastically accelerate forward hybridization kinetics, are available. Further, many of these preparations contain a surfactant that prevents the probe from binding to the filter, thereby minimizing background hybridization. While somewhat costly and of proprietary formulation, these hybridization reagents are convenient and work remarkably well. Plus, elimination of formamide from the hybridization recipe is an added bonus.

Urea

The rather toxic properties of formamide, a common component of traditional prehybridization and hybridization buffers, clearly present intrinsic exposure dangers in the lab. Alternative formulations have been developed containing urea; hybridization solutions containing 6M urea are equivalent to those

³100× Denhardt's solution = 2% Ficoll 400, 2% polyvinylpyrrolidone, 2% BSA (fraction V).

containing 50% formamide with respect to lowering T_m . In particular, the ECL™ nucleic acid labeling and detection system (GE Healthcare) is designed in this manner. This system is compatible with the common nucleic acid blotting applications and tends to produce posthybridization detection images that are very low in background.

Urea-based hybridization formulations are compatible with both DNA probes and RNA probes (Simard et al., 2001). The concentration of urea needed to be effectively destabilizing ranges from 2–6 M. Urea stock solutions should be freshly prepared prior to use.

Hybridization temperature

In most applications, nucleic acid probes that are at least several hundred bases long can be hybridized at 42° in a generic hybridization solution which, among other things, often consists of 50% formamide or a urea equivalent. When more precisely defined hybridization conditions are required or when empirical determination must be made, one may first derive the T_m of the duplex and conduct hybridization at a suitable temperature. With respect to the following equations, the contribution of each of the salient parameters has been discussed previously.

T_m for long probes

The thermodynamic stability is described as follows:

Double-stranded DNA (DNA:DNA) molecules (Bolton and McCarthy, 1962):

$$T_m = 81.5^\circ + 16.6 \log[\text{Na}^+] + 41(\%G + C) - 0.63(\% \text{ formamide}) - (500/L)$$

DNA:RNA hybrid molecules (Casey and Davidson, 1977):

$$T_m = 79.8^\circ + 18.5 \log[\text{Na}^+] + 58.4(\%G + C) + 11.8(\%G + C)^2 - 0.5(\% \text{ formamide}) - (820/L)$$

Double-stranded RNA (RNA:RNA) molecules (Bodkin and Knudson, 1985):

$$T_m = 79.8^\circ + 18.5 \log[\text{Na}^+] + 58.4(\%G + C) + 11.8(\%G + C)^2 - 0.35(\% \text{ formamide}) - (820/L)$$

where:

T_m = melting temperature of duplex molecules.

(% G+C) = percentage of guanine + cytosine content, expressed as a mole fraction.

$[\text{Na}^+]$ = log of the sodium concentration, expressed in molarity.

L = number of bases that participate in the actual hybridization.

T_m for oligonucleotide probes

For shorter probes such as oligonucleotides (Chapter 12), the T_m of the oligomer can be quickly estimated by assigning 2° for each adenine and thymine and 4° for each guanine and cytosine (Itakura et al., 1984):

$$T_m = 4^\circ(G + C) + 2(A + T)$$

It is important to realize that this relationship holds only for short oligonucleotides (11–27 bases long) and is based on hybridization in 1M Na⁺ (6× SSC), in the complete absence of any organic solvents. By definition, at the T_m only 50% of all duplexes are stable; therefore, hybridization is often performed 3–5° below the calculated T_m for short oligonucleotides to promote complete target strand saturation. At $T_m - 5^\circ$, all exact match duplexes will form. It is important to realize that in the design of an oligonucleotide probe there may have been certain ambiguities pertaining to the sequence of the probe. If the oligonucleotide probe is not an exact match, then it will be necessary to lower the stringency of hybridization to tolerate mismatches. If more than one mismatch is involved, it will be necessary to lower the temperature of hybridization even more.

When longer oligonucleotides are involved (14–70 bases), the T_m can be calculated (Sambrook and Russell, 2001) by

$$T_m = 81.5 - 16.6(\log[\text{Na}^+]) + 41(\%GC) - (600/N)$$

where N = the length of the oligonucleotide (assuming a perfect match between probe and target).

These same basic principles apply to the design of oligonucleotides that will act as primers for PCR (Chapter 18). It is worth noting that there are many web sites which allow investigators to determine the exact T_m of oligonucleotides prior to ordering them so as to match the probe as closely as possible to one's experimental parameters.

Hybridization and the Northern analysis

The attempted hybridization between nucleic acid probe and target sequences consists of three major components: prehybridization (also known as blocking), hybridization, and posthybridization washing (also known as the stringency washes). The detection of hybridization events, the subject of the next chapter, is a different art entirely.

The exact hybridization conditions for many of the universally available probes have been defined in the literature. Many generic types of hybridization buffers and conditions have been recommended, combinations that purportedly yield the “perfect blot every time”. Although it is true that the kinetic behavior

of some families of nucleic acid molecules can be predicted with a fair degree of accuracy, knowledge of the parameters that directly influence hybrid formation is invaluable in fine-tuning the conditions of hybridization. This becomes a paramount issue when using probes that are not exactly complementary to the target(s) (exact match) or perhaps when using a nucleic acid probe to screen blots or libraries in phylogenetic studies. It should be noted that temperature and salt concentration are the two variables most often adjusted in order to modify the specificity of hybridization, when such modifications become necessary.

Prehybridization: filter preparation

As described in Chapter 11, nucleic acid samples should be fixed to filter membranes immediately after transfer. The filters can then be stored for extended periods. Before hybridization, filters are subjected to a prehybridization incubation (1) to equilibrate the filter in a buffer identical or very similar to the one to be used for the actual hybridization; and (2) to block the filter paper completely. Blocking (Meinkoth and Wahl, 1984) means covering any part of the filter that is not already occupied by experimental RNA samples with sheared, denatured (heterologous) DNA from a species unrelated to the biological origin of the RNA; usually salmon sperm DNA or calf thymus DNA is selected⁴. Alternatively, newer methods involve blocking filters, especially nylon, with casein or other protein, rather than with heterologous DNA. In general, protein blocking requires less time than blocking a filter with heterologous DNA. It is ideal for isotopic and non-isotopic detection systems alike, yielding an outstanding signal-to-noise ratio with very low background. Thorough prehybridization is important to ensure that, when added, the probe molecules do not stick in a non-specific fashion over the surface of the filter paper, which would cause unacceptable high background levels.

Numerous recipes have been concocted for prehybridization, their exact composition being a function of the type of filter and probe used in a specific experimental application. The recommendations contained herein have been used with great success in this and other laboratories.

Protocol: prehybridization for long probes

1. Pre-wet the filter in $5\times$ SSC or $5\times$ SSPE for 2 to 3 min.

Note: SSPE is a better buffer than SSC, especially in the presence of formamide (Meinkoth and Wahl, 1984), because the phosphate in SSPE actually mimics the nucleic acid phosphodiester backbone and helps to block the filter and improve the background

⁴Heterologous DNA, used as a carrier and a filter membrane block is prepared as a 10 mg/ml stock solution in water. The DNA is boiled to facilitate dissolving and then sheared by drawing it repeatedly through an 18-gauge needle. Stock solutions of heterologous DNA are maintained at 4° for short-term storage or at -20° for long-term storage. Just prior to use, an appropriate aliquot must be denatured again by boiling for 10 min.

at the detection stage. This strategy is especially desirable when experimental RNA is immobilized on nylon (+) filters, due to the natural affinity of the membrane for the negatively charged phosphodiester backbone of the nucleic probe. $20 \times$ SSPE = 3 M NaCl, 200 mM NaH_2PO_4 , 20 mM EDTA, adjust pH to 7.4. $20 \times$ SSC = 3 M NaCl, 0.3 M Na_3 -citrate, adjust pH to 7.0.

2. Prehybridize the filter(s) for 3 to 4 hours at 42° in an excess volume (250–1000 $\mu\text{l}/\text{cm}^2$) of prehybridization buffer consisting of $5 \times$ SSPE, $5 \times$ Denhardt's solution, 0.1% SDS, 100 $\mu\text{g}/\text{ml}$ freshly denatured salmon sperm DNA, 50% deionized formamide. The following is the proper method for assembling the components of the prehybridization buffer:
 - a. Prewarm freshly deionized formamide to 45° .
 - b. In a separate bottle or tube, mix SSPE, Denhardt's solution, and SDS and place at 45° . The SDS will go into solution when the remaining ingredients are added.
 - c. In a fresh polypropylene or glass test tube, mix the required salmon sperm DNA in the volume of sterile water necessary to bring the prehybridization buffer up to final volume. Boil the mixture for 5 to 7 min.
 - d. Carefully add the boiled salmon sperm DNA to the SSPE, Denhardt's, SDS mixture, and then add the entire mixture directly to the prewarmed formamide.

Swirl carefully to mix. Store this hybridization solution at 42° until ready to use. The reason for assembling the buffer in this fashion is to eliminate renaturation of denatured salmon sperm DNA that could occur by using the more common strategy of plunging a small volume of concentrated stock DNA on ice. Immediate addition of formamide effectively reduces the rate of renaturation, especially when all reagents are prewarmed. After mixing, the buffer is ready to be used immediately. Typically, prepare 100 ml of this buffer to prehybridize two blots and reserve 25 ml, which can be used for hybridization.

Note 1: Routine prehybridization and hybridization are most expediently carried out in a plastic container with a perfectly flat bottom, whose surface area is only slightly greater than the area of the filter. Several configurations of Rubbermaid® food storage boxes and Nalgene® utility boxes work well. Be sure to cover the container with a securely fitting lid to prevent evaporation. This approach precludes introducing air bubbles into heat-sealable plastic bags and avoids spillage of high specific activity hybridization buffer at the conclusion of the hybridization period.

Note 2: Placing the hybridization vessel on a gently moving (30–50 rpm orbital shaker incubator (e.g., Lab-Line)), will circulate the buffer evenly over the entire surface of the filter and yield superior data.

Note 3: Prehybridization with a large (excess) volume of buffer usually favors very low levels of non-specific interaction between the probe (when added) and the filter. After blocking, the volume of the hybridization buffer can be reduced before adding the probe so as to maintain the probe at a respectable concentration.

3. At the conclusion of the prehybridization period pour off the prehybridization buffer and add a fresh aliquot (optimally 100 $\mu\text{l}/\text{cm}^2$) of hybridization buffer. In most cases, the prehybridization and hybridization buffers will be identical (see step 2 for preparation).

Probe denaturation

Double-stranded probes must be denatured prior to use. Denaturation is often accomplished by boiling or by raising the pH. Because of the enormous variation

in labeling approaches in non-isotopic systems, it is best to follow the instructions of the manufacturer for probe denaturation. In the case of direct enzyme labeling, the probe will already be single-stranded, and can be used directly without any further manipulation.

CAUTION: Radiolabeled probes should not be boiled to be denatured because of the dangers associated with radioactive steam. Instead, add 0.1 volume 1 N NaOH to the probe (which is usually in a volume of 50–100 μ l) and incubate at 37° for 10 min or at room temperature for 30 min. The probe can then be added directly to the hybridization solution, without prior neutralization, as long as the total probe volume is less than 150 μ l (assumes a minimum of 5 ml hybridization buffer).

Hybridization

4. Add the denatured probe to the buffer and hybridize for 12 to 16 h (overnight) at 42°. Make certain that the cover of the hybridization chamber is securely in place.

Note 1: For the detection of rare mRNAs or single copy genes, use 5 to 20 ng of probe per milliliter of hybridization buffer or $1\text{--}5 \times 10^6$ cpm/ml.

Note 2: If the mass of the probe is low, increasing the concentration of the salmon sperm DNA will improve hybridization kinetics.

Note 3: A total of 25 ng probe that will base-pair with the molecular weight standards may also be added to produce a visual record of the standards along with experimental samples on the same film.

Posthybridization stringency washes

Depending on the nature of the probe and the stability of the hybrid, the investigator may need to adjust the exact posthybridization washing conditions for each probe. It is very important to keep in mind that in the context of posthybridization washes, stringency is a function of ionic strength, temperature, and time; thus, it may not be suitable to cut down on the washing time by increasing the temperature and/or lowering the salt.

5. At the conclusion of the hybridization period, pour off the hybridization buffer⁵ and wash the filters to remove all of the probe molecules that did not participate in the formation of hybrids.

a. Wash filters for 30 s, in a room temperature solution of 2 \times SSPE, 0.1% SDS to remove most of the probe in the box and remaining on the filter membrane.

b. Wash filters twice for 15 min, in 200 to 300 ml of a room temperature solution of 2 \times SSPE, 0.1% SDS.

c. Wash filters twice for 15 min, in 200–300 ml of a solution of 0.1 \times SSPE, 0.1% SDS at 37°.

Note: If the background is too high, a final wash in 0.1 \times SSPE, 0.1% SDS at 42° to 50° may be useful. It is likely that the optimal wash conditions will need to be determined empirically for each new probe.

⁵Dispose of used hybridization and wash solutions appropriately, e.g., as radioactive waste.

6. Rinse the filter very briefly in $2\times$ SSPE and place on a piece of Whatman paper just long enough to blot excess buffer from surface of the filter.

Important: *do not allow the filter to dry out to any extent, especially when chemiluminescence is the method chosen for detection.*

- 7a. If using chemiluminescence or chromogenic detection technology, proceed from this step directly to the detection buffers, according to the directions of the manufacturer of the system.
- 7b. If using isotope-labeled probes, wrap the filter in plastic wrap and set up autoradiography (Chapter 14).

Note: Be sure to keep filters damp if planning to remove the probe after autoradiography in order to hybridize to another probe(s).

Probe removal and rehybridization

Filters with immobilized RNA or DNA sequences can be screened repeatedly with several different probes if the filter is handled and stored properly. This applies to Southern blots, Northern blots, dot-blot, colony lifts, and plaque lifts. Rehybridization is often desirable because (1) the preparation of the blot may represent a lot of effort; (2) there may be limited or no sample remaining with which to prepare another blot; (3) it is important to probe a single filter for the expression of housekeeping (control) genes in addition to screening for the modulation of genes of interest; (4) in the case of dubious results or high background, re-screening the same filter with the same probe may clear up ambiguity; and (5) one may decide later on to rescreen a blot based on new information or the availability of new probes. Consequently, proper handling of the filter is important.

Prior to the first hybridization, the filter can be stored dry and out of direct sunlight. In this lab, filters that have been UV crosslinked are stored dry between two sheets of Whatman 3 MM paper in the drawer set aside exclusively for filter storage. Placing a notebook or other light weight on top of the filter(s) helps to keep them flat. Filters can be stored this way for months, if necessary.

Immediately following hybridization detection (Chapter 14), filters are subject to high stringency washes to remove the old probe. In the event that probe removal is not convenient right after probe detection, the filter should remain wrapped in plastic to keep it damp. If the filter does dry out prior to probe removal, it will be all but impossible to remove the probe later on. Some investigators don't bother to remove previously hybridized probe, claiming that it does not interfere with subsequent characterizations of the blot. The truth is that old probe may or may not interfere with subsequent hybridization and analysis of the data. Table 13.2 is structured to help one decide if probe removal is necessary prior to rehybridization. As a general rule, it is worth the investment of time to removal old probe prior to rehybridization or long-term storage. Deciding later on to try to remove a previously hybridized probe is often a decision made too late.

Table 13.2 Probe Removal Matrix

		Advantages	Disadvantages
Radioactive probe	Remove probe	1. No signal expected from previously hybridized probes 2. Can use the same probe again for confirmation of previous hybridization data, as needed	Investment of time and reagents
	Retain probe	No additional processing necessary	1. Previous probe likely to produce bands 2. Unable to hybridize to the same target sequences again
Non-isotopic probe	Remove probe	Can use the same probe again for confirmation of previous hybridization data, as needed	Investment of time and reagents
	Retain probe	No additional processing necessary	1. Residual detection chemistries may contribute to high levels of background in subsequent detections 2. Unable to hybridize to the same target sequences again

Protocol: generic method for probe removal

Each type of filter has a recommended probe removal protocol. The protocol presented here is generic enough to work well with most nylon filters. Filters should be kept moist following the detection procedure until the following steps can be performed to remove the old probe.

1. Wearing gloves, carefully remove damp filters from protective plastic wrap. If the previously hybridized probe was radioactive, be sure to observe precautions for the safe handling of the filter.
2. Immediately immerse the filter in an excess volume of 50% formamide, 0.5× SSPE which has been preheated to 65°. **CAUTION:** Be sure to tightly seal the container in which this wash is being performed to prevent exposure to formamide vapors.
3. Incubate for 45–60min on an orbital platform at 50rpm, and maintain the temperature throughout. If an orbital platform is not available, then the container in which the wash is being performed should be agitated manually every 10 min.
4. Carefully transfer the filter to a clean container and rinse in two changes of 0.1× SSPE for 2 min each with gentle agitation.

5. Place the nylon filter, face up, on a sheet of Whatman 3MM paper and blot the excess $0.1 \times$ SSPE.
6. If a radioactive probe was used previously, it is a good idea to keep the newly washed filter damp, place it in plastic wrap and expose it to a sheet of X-ray film to ensure that all of the previously hybridized probe was removed. The likelihood of probe removal can be estimated by passing a Geiger counter over the surface of the film to assess the amount of residual radioactivity.

Note: If bands do appear on the X-ray film, repeat steps 1–5. If the probe fails to strip off the filter after a second attempt, the probe will probably remain on the filter. In such an event, one could store the filter until the probe goes through several half-lives, after which the filter can be hybridized to a different probe sequence.

References

- Blake, R.D. and S.G. Delcourt. (1996). Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.* **24**(11), 2095–2103.
- Bodkin, D.K. and D.L. Knudson. (1985). Assessment of sequence relatedness of double-stranded RNA genes by RNA–RNA blot hybridization. *J. Virol. Methods* **10**, 45–52.
- Bolton, E.T. and B.J. McCarthy. (1962). A general method for the isolation of RNA complementary to DNA. *Proc. Natl. Acad. Sci. USA* **48**, 1390–1397.
- Bonner, T.I., D.J. Brenner, B.R. Nefeld and R.J. Britten. (1973). Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.* **81**, 123–135.
- Britten, R.J. and E.H. Davidson. (1985). Hybridization strategy. In *Nucleic Acid Hybridization: A Practical Approach* (B.D. Hames and S.J. Higgins, Eds). IRL Press, Washington, DC. pp. 3–16.
- Casey, J. and N. Davidson. (1977). Rates of formation and thermal stabilities of RNA: DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4**, 1539–1552.
- Chang, C.T., T.C. Hain, J.R. Hutton and J.G. Wetmur. (1974). Effect of microscopic and macroscopic viscosity on the rate of renaturation of DNA. *Biopolymers* **13**, 1847–1858.
- Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23**, 641–646.
- Itakura, K., J.J. Rossi and R. Wallace. (1984). Synthesis and use of synthetic oligonucleotides. *Annu. Rev. Biochem.* **53**, 323–356.
- Lederman, L.L. (1981). The rate of nucleic acid annealing to cytological preparations is increased in the presence of dextran sulfate. *Analyt. Biochem.* **117**(1), 158–163.
- Marmur, J. and P. Doty. (1961). Thermal renaturation of deoxyribonucleic acids. *J. Mol. Biol.* **3**, 585–594.
- Marmur, J. and P. Doty. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.* **5**, 109–118.
- Meinkoth, J. and G. Wahl. (1984). Hybridization of nucleic acids immobilized on solid supports. *Analyt. Biochem.* **138**, 267–284.
- Sambrook, J. and D.W. Russell (Eds). (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, New York.
- Simard, C., R. Lemieux and S. Côté. (2001). Urea substitutes toxic formamide as destabilizing agent in nucleic acid hybridizations with RNA probes. *Electrophoresis* **22**, 2670–2683.

-
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.
- Subirana, J.A. and P. Doty. (1966). Kinetics of renaturation of denatured DNA. I. Spectrophotometric results. *Biopolymers* **4**, 171–187.
- Thrower, K.J. and A.R. Peacocke. (1968). Kinetic and spectrophotometric studies on the renaturation of deoxyribonucleic acid. *Biochem. J.* **109**, 543–557.
- Wahl, G.M., M. Stern and G.M. Stark. (1979). Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. *Proc. Natl. Acad. Sci.* **76**, 3683–3687.

This page intentionally left blank

14 Principles of Detection

Rationale	301	Protocol: autoradiography	310
Autoradiography	302	Non-isotopic procedures	312
Handling of filter membranes	304	Biotin	313
X-ray film	305	Digoxigenin	313
Safelight	306	Fluorescence imaging	314
Exposure time	306	Direct enzyme labeling	314
Intensifying screens	307	Detection by chemiluminescence	314
Fluorography	308	Substrate for chemiluminescence	315
Preflashing film	309	Chromogenic detection procedures	317
Type of cassette	309	Digital imaging systems	318
Development and fixation	310	References	320

Rationale

The outcome of an experiment is only as reliable as each step involved in the generation of data. The selection of labeling and detection techniques is governed primarily by the required degree of sensitivity and resolution and is also influenced by the format of the experiment: filter-based, solution hybridization, *in situ* analysis, and so forth. The two goals at the end of every experiment involving hybridization are to (1) **Localize** the probe, because it is necessary to determine where on the filter paper, or elsewhere in the assay, the probe managed to base-pair stably to its target; and (2) **Quantify** the amount of the probe that has been retained in each band or is otherwise hybridized to target molecules found within a particular region of a tissue section.

Broadly, there are two standard detection strategies: those involving radioisotopes, and non-isotopic techniques, each of which are discussed in this chapter. The non-isotopic methods are becoming increasingly sophisticated, accommodating detection on filter membranes, in microtiter plate format, directly from gels, and by direct video capture, to list but a few of the choices. Remarkable refinements in non-isotopic systems allow the investigator to realistically question the use of radioactivity at all. Isotopic methods, obviously, involve radiolabeling, coupled with detection by autoradiography alone or in conjunction with scintillation counting.

Autoradiography¹

Autoradiography is a simple and sensitive photochemical technique used to record the spatial distribution of radiolabeled compounds within a specimen or an object. In the context of nucleic acid-related study, radiolabeled probes are hybridized to target sequences that are immobilized on a solid support, in solution, or *in situ*. The capture of ionizing radiation and photons by an emulsion placed in direct contact with a radioactive source provides a relatively permanent record of the decay events associated with unstable radionuclides. Autoradiography is subdivided into two broad groups, commonly referred to as macroautoradiography and microautoradiography, indicative of the type of specimen containing the radioactivity, the type of emulsion necessary for image formation, and the method of examining the results.

The most common applications of autoradiography in molecular biology include the quantitative analysis of nucleic acid hybridization events. If used in conjunction with electrophoresis, the resultant imagery will provide a qualitative aspect as well. For example, in Northern analysis (macroautoradiography), the extent of hybridization and probe specificity to RNA target molecules is assessed by the magnitude and location of radioisotope emissions from the filter membrane. Microautoradiography, on the other hand, is performed by coating a thin section of a larger specimen with a light-sensitive emulsion. This technique, also known as *in situ* autoradiography, is usually performed on a microscope slide. Because this technique has the advantage of maintaining natural cellular geometry, the intracellular distribution of radiolabel can be easily determined by light or electron microscopy (i.e., by counting the grains).

In macroautoradiography, a piece of high-speed X-ray film is placed in direct contact with a plastic-wrapped filter membrane. This is commonly known as direct exposure. In some applications, such as the nuclease protection assays, isotopes can be detected directly in polyacrylamide gels without the need for blotting onto a filter. The film base is a flexible piece of polyester [Kodak films used for autoradiography are 0.007-inch (7 mil) thick] that has been coated on one or both sides with an emulsion. The radio- and light-sensitive emulsion itself consists partly of sensitive silver halide (bromides, chlorides, or mixed halide) crystals, to which energy is released by ionizing radiation. During film exposure, energy is absorbed in the silver-halide grains within the film emulsion, with the release of electrons. The resulting negatively charged "sensitivity specks", i.e., faults in the crystal lattice, attract positively charged silver ions, thereby forming an atom of metallic silver.

Exposure of X-ray film generates the so-called latent image, an invisible precursor to the visible image manifested when the film is immersed in a suitable photographic developing solution. The latent image, consisting only of developable silver grains, is not visible in the darkroom (under safelight conditions) and the emulsion remains exquisitely light-sensitive. Formation of the latent

¹Adapted, in part, from Eastman Kodak Company.

image can be improved by preflashing the film (described below) and by low temperature exposure. Once formed, the latent image is not stable for long exposure periods; potential latent image instability is exacerbated by the presence of oxygen and moisture on the film. The final developed image may be assessed by visual inspection or, with greater precision, by digital image analysis. The overall developing process is outlined in Fig. 14.1. There are distinct stages in the processing of the film: development, washing, fixation, and finally drying; these stages can be performed manually or with an automatic film processor. In either case, it is wise to standardize the developing parameters, particularly when perform these steps manually, so as to allow meaningful comparison from one film to the next.

When immersed in a photographic developer, the exposed silver halide grains in the emulsion of the film are reduced to metallic silver, greatly amplifying the latent image. The amount of time required to complete the developing process is variable, depending on the degree of exposure of the film, the temperature of

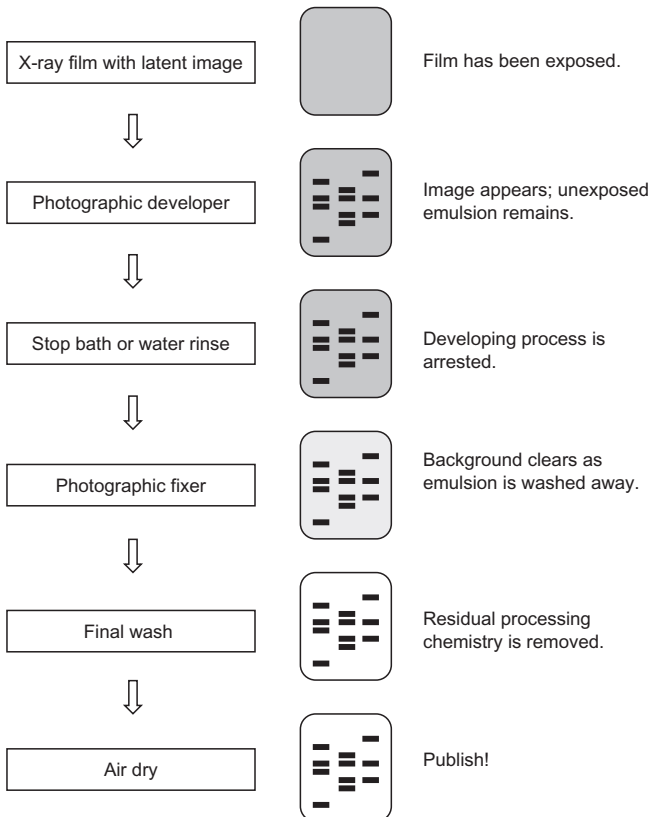


Figure 14.1 Developing scheme for X-ray film. Although the choice of film is dependent on the application (autoradiography vs. chemiluminescence), the mechanics of X-ray film processing are identical.

the developer, the age of the developer, and the tolerance of the investigator for background. In this lab Kodak D-19 and D-76 developer are used routinely.

The action of the developer is arrested by immersion of the film in a chemical "stop bath". A variety of stop bath solutions may be purchased, though a homemade version can be concocted by preparing a solution of 1% acetic acid. Arresting the activity of the developer at precisely the right time is very important because over-development can obscure any image that may have been captured on film. Keep in mind that a piece of X-ray film withdrawn from the developer tray is still coated with the developer, and that the developing process will continue until the action of the developer has been neutralized. In the absence of a suitable stop bath formulation, the film should be immersed quickly in water to dilute and wash off the developer. These actions will not only arrest the developing process but also extend the life of the fixer. Next, the film is immersed in a chemical fixer solution (e.g., standard Kodak fixer), which dissolves away the unexposed silver halide, producing the characteristic X-ray film appearance. A thorough, room-temperature washing of the film in water, after the fixation step, is needed to remove processing byproducts. Once the film has been completely dried, a relatively stable image is in place.

When developing films manually, it is necessary to develop, wash, and fix each film for the same length of time. Without this type of quality control it may be difficult to determine the extent to which observed data variations are due to experimental manipulation or to variation in the mechanics of developing the film. One of the advantages of manual film processing is that the investigator can extend the developing time in order to accentuate faint bands, or reduce the developing time in order to minimize background or the over-development of extremely strong signals from one or more bands. Machine-developed autoradiograms have the advantage of being developed according to a standardized process; however, one is at the mercy of the machine.

The mechanics of an autoradiographic exposure (reviewed by Bonner, 1987) and the choice of emulsion are dependent on the energy emission spectrum of the isotope, the image quality requirements, and the anticipated amount of radioactivity in the system. There are no hard and fast rules about exposure time and exposure strategies; these parameters must be empirically determined with each experiment.

Handling of filter membranes

Filter membranes used for Northern analysis, dot-blotting, and similar techniques should not be allowed to dry out completely after the last stringency wash. Instead, these filters should be blotted briefly by resting them on a piece of Whatman 3 MM paper, radioactive side up, and then wrapped in plastic. Maintaining filter membranes in a damp state will greatly facilitate the post-autoradiography removal of the hybridized probe by an extremely high-stringency wash. Thus, a filter membrane can be screened several times with other probes of interest. Removal of probe from membranes that are allowed to dry completely

is significantly more difficult than removal from damp membranes, if not altogether impossible. *Never* place a damp or wet filter membrane in direct contact with a sheet of X-ray film. The two will form a permanent symbiosis of sorts, resulting in the ruination of both filter and film.

X-ray film

The choice of X-ray film is dependent on whether high-energy or low-energy emitters were used as label and the required level of resolution. BioMax film² (Kodak) and Hyperfilm (GE Life Science) are the latest generation of X-ray films for autoradiography and chemiluminescence. Different formats of these high performance films are available to accommodate high resolution applications, high sensitivity applications, and general purpose use. For the convenience of the reader, Table 14.1 suggests the proper film for autoradiography and chemiluminescence in the context of the required level of sensitivity and resolution.

The maximum resolution films are recommended for use with the lower energy β -emitters ³⁵S, ³³P, and ¹⁴C, while the high performance and multipurpose formats can be used to detect ³²P, ³³P, ¹²⁵I, and ³⁵S as well as applications involving chemiluminescence detection. The newer films have the distinct advantage of being clear base films, rather than having the traditional gray tint associated with older film formulations. The clear base allows the visualization

Table 14.1 Select X-ray Films for Autoradiography and Chemiluminescence

	Autoradiography	Chemiluminescence
Maximum sensitivity	BioMax MS, Hyperfilm MP	BioMax Light, Hyperfilm ECL
All purpose usage	BioMax MS, BioMax XAR, Hyperfilm MP	BioMax Light, BioMax XAR, Hyperfilm ECL
Maximum resolution	Hyperfilm MP BioMax MR	BioMax Light Hyperfilm ECL

Film selection for hybridization detection is based on the requirement for sensitivity and resolution. Biomax MS[®], BioMax Light, and Hyperfilm ECL are double-emulsion film, while BioMax MR is a single-emulsion film. Hyperfilm MP is a double-emulsion, all purpose film, as is BioMax XAR. X-Omat LS is a double-emulsion, lower sensitivity blue tint film that requires longer (2×) exposure times. Most X-ray films are available in two configurations, either in alternate interleaved packing that protects each sheet of film during shipment and prevents static electricity from exposing the film as it is withdrawn from the box, or packaged such that each sheet of film is individually wrapped in a light-tight envelop (e.g., ReadyPack format). MS = maximum sensitivity; MR = maximum resolution; MP = multipurpose; and ECL = enhanced chemiluminescence.

²The time-honored Kodak X-Omat XAR film family was reformulated and is now known as the BioMax AR film series and is designed specifically for life science applications.

of even the faintest bands on the film and is quite compatible with the contemporary methods for scanning and digital image analysis.

Unopened packages of X-ray film should be stored between 50° and 70°F and shielded from penetrating radiation. Once opened, storage between 30–50% relative humidity is recommended. For extended periods, boxes of film should be stored upright, rather than stacked in a horizontal position, so that the film does not crease or buckle, which could result in background or uneven processing.

It is very important to protect the undeveloped film from any stray light. While almost any room can be turned into a photographic darkroom rather economically, it is of great importance to ensure that no light enters the darkroom from under the door(s), around the edges of windows, through openings in the walls or ceiling, or even from indicator lights on equipment that might be in the room, including computer monitors.

Safelight

Undeveloped X-ray films remain light sensitive until processed through the photographic fixer solution, though limited low-intensity illumination is permissible to allow movement and efficient handling in the darkroom. A safelight typically consists of an ordinary 15-watt frosted light bulb that shines through a filter, the color of which is compatible with the film being used. The most commonly used films for chemiluminescence and autoradiography are blue-green sensitive, so the recommended darkroom setup features a red GBX-2 safe light filter³. The term safelight is something of a misnomer, though, because prolonged exposure to safelight illumination will fog the film. Ideally, the film should never come any closer to the safelight than four feet. Be aware also that safelight filters fade over time and with extended use, so testing the filter periodically is recommended. This can be performed easily by first developing an unexposed sheet of X-ray film with the safelight on and then developing another unexposed sheet of film with the safelight off. Any fogging of the first film suggests that the safelight filter needs to be replaced.

Exposure time

There is no limit to the number of exposures that can be made from a single radioactive source other than the half-life of the material itself. In the case of high-energy emitters, a Geiger counter probe placed in the vicinity of the filter is good way to assess the activity on a filter membrane. Obviously, a source manifesting 3000 cpm will require a much shorter exposure period than a source registering but a few unimpressive “clicks” above background. The lack of obvious signal should not be an immediate discouragement; with autoradiography, it is

³An excellent economical safelight for quickly setting up a darkroom is the GBX-2 Safelight (Kodak No. 141 6627). It is small, portable, and durable.

not unusual to have to wait three days or longer to observe hybridization signals corresponding to single-copy genes or rare mRNAs.

It is often useful to do an overnight exposure at room temperature to assess the extent of background and non-specific filter hybridization; subsequent exposure tactics are based on the image produced from this test exposure. In the case of low-activity filters one can reduce the exposure time with the use of an intensifying screen, which is a temperature-dependent process. Low temperature exposure alone, however, will not improve the direct exposure image formation process.

Intensifying screens

Autoradiographic detection of ionizing radiation can be improved through the use of intensifying screens, which consist of a flexible polyester base coated with luminescent phosphor crystals. X-rays striking the phosphor coating generate visible light which can further expose a sheet of X-ray film. Thus, intensifying screens are commonly used to lessen the time, and the radiation, needed to produce an image on the X-ray film within the exposure cassette.

The popular intensifying screens once in widespread use for autoradiography were calcium tungstate (CaWO_4) phosphor-coated screens that produce blue light upon excitation. The newer generation of intensifying screens (e.g., DuPont's Cronex Quanta III Intensifying Screen) are known as rare earth screens and are much faster than the older style CaWO_4 -coated screens. Rare earth screens are coated with gadolinium oxysulfide, lanthanum oxybromide, or yttrium-oxysulfide and, depending on the phosphor, produce a bright green or blue glow when excited. The rare earth screens exhibit very low-level noise (background signal), and are much more efficient in the conversion of X-rays into light; as such they are nicely compatible with standard autoradiography procedures. Thus, the rare earth screens are ideal for use with exposures that would otherwise require longer than one day. If used, CaWO_4 intensifying screens should be stored in the dark for several hours or overnight before setting up the exposure is also worthwhile because "after glow" effects from non-dark-adapted screens can fog the film; however, dark-adaptation is not necessary with rare earth screens.

Unlike double-emulsion X-ray film, which has no sidedness, only one side of an intensifying screen is phosphor-coated, so proper orientation is essential. In this lab, to minimize confusion and possible error in the dark room, strips of heavy tape have been placed along the edges of the *non-functional* side of the intensifying screen so that feeling the tape on one side will help the user properly position the phosphor-coated side facing the film. The proper juxtaposition of intensifying screen and X-ray film within the cassette is also important if the screen is to be of any use whatsoever during the exposure period. The film is placed between the specimen and intensifying screen so that the film can "see" scintillation events generated by the screen, in addition to being exposed directly by the ionizing radiation associated with isotope decay.

The proper location of a second intensifying screen (if used) is behind the specimen. Placement of the film between two intensifying screens with the specimen outside the “sandwich” is not recommended because it will reduce image sharpness (it may improve sensitivity, however). One may observe a minor decrease in image sharpness even if a single intensifying screen is used, due to the spread of β particles from their source as they travel through the film to the screen in addition to the associated emission of light. In essence, there is considerable “bouncing around” of β particles within the cassette as they strike the intensifying screen. The additional spread of the energy causes a larger area of the film to be exposed than when only direct exposure to the β energy source is carried out. This is the cost of the sensitivity enhancement from the use of an intensifying screen.

Autoradiography with intensifying screens is performed at -70° because (1) film response to low light intensities is improved dramatically at low temperatures; and (2) intensifying screens perform best when cooled to -70° . The net result, then, is a greater degree of exposure per decay event than in the absence of an intensifying screen. Intensifying screens improve the detection of the high-energy β emitters ^{45}Ca and ^{32}P and γ ray emitters such as ^{125}I (Laskey and Mills, 1977; Swanson and Shank, 1978). Traditional intensifying screens are not effective with the low energy emitters ^3H , ^{35}S , or ^{14}C . In the case of ^{33}P , used mostly for old-style DNA sequencing (Zagursky *et al.* 1991), CaWO_4 screens were ineffective because the autoradiographic characteristics of ^{33}P are similar to those of ^{35}S . Some of the newer rare earth screens have been formulated specifically for use with low energy emitters, too.

Finally, it is very important that the sensitivity of the film used matches the light emission spectrum of the intensifying screen. The luminescent phosphor composition of a particular rare-earth intensifying screen will result in the emission of light at a characteristic wavelength, so it is important to ensure that the film (blue or green-sensitive) is compatible with the screen. This information may be found on the product insert that accompanies the screen or by first calling the manufacturers' technical services department prior to ordering.

Fluorography

Fluorography is an intensification process by which some of the energy associated with isotope decay is converted to light by the interaction of radio-decay particles with a compound known as a fluor, which then exposes the X-ray film. As with intensifying screens, fluorography is an intensification procedure, though it is used mainly when low-energy emitters, especially ^3H , are employed as radiolabels (Bonner and Laskey, 1974; Randerath, 1970). Fluorography is accomplished through the impregnation of a specimen, such as a polyacrylamide gel, with a fluorographic solution (a scintillant or “scintillator”), after which the gel is dried down for efficient autoradiographic detection. Commercially available fluorographic formulations include mixtures of 2,5-diphenyloxazole, also known as PPO, or sodium salicylate. While fluorographic

impregnation does not improve detection performance with ^{32}P or ^{125}I , it may enhance detection of ^{14}C , ^{33}P , and ^{35}S , and is almost always used with ^3H . In this way, fluorographic detection differs from signal amplification through the use of intensifying screens, which may be viewed as a type of external fluor. Fluor-impregnated samples should definitely be exposed to preflashed film (Laskey and Mills, 1975) at -80° , to ensure that the response of the film is linear. Fluorography is not a commonly used methodology in conjunction with most routine molecular biology applications.

Preflashing film

To increase the sensitivity of film to low intensities of light and thus increase the linear response of the film, an investigator may elect to preflash (hypersensitize) a sheet of film just prior to setting up an autographic or fluorographic exposure. By providing a short, controlled burst of light, the film essentially becomes charged or primed, meaning that the latent image formation process is initiated. The latent image formation process is then completed by light photons from fluorographic exposure or from intensifying screens. Priming the film in this manner overcomes the rather narrow linear response of the film, thereby greatly enhancing its sensitivity. Calibrated pre-flash units can be purchased from a number of suppliers, and the exact conditions for use are dependent on local darkroom conditions. In the absence of an intensifying screen, preflashing is not an effective means of enhancing latent image formation by ionizing radiation. Preflashing is useful for increasing the sensitivity of X-ray film when used in conjunction with chemiluminescence, though it is not routinely performed in this lab.

Type of cassette

The selection of an appropriate light-tight apparatus to accommodate autoradiographic exposures is probably the least involved decision of all. There is an exposure apparatus to fit every budget, all of which can be used both for autoradiography and chemiluminescence. The least expensive style is the vinyl-covered exposure cassette. This type of exposure holder is best sandwiched between two 1 cm-thick sheets of Plexiglas to block high-energy β particles which would otherwise pose a potential health hazard to the investigator and others in the vicinity. The more commonly used models are constructed from lightweight, vinyl-covered aluminum and include permanently mounted intensifying screens, as are stainless-steel X-ray film holders. Within these more deluxe cassettes, polyurethane pads help to maintain continuous uniform contact between the film and the filter. When assembling the exposure apparatus, it is of paramount importance to immobilize a filter membrane; even the slightest movement of the filter at any time during the exposure period will produce a fuzzy image. Immobilization by mounting the plastic-wrapped filter on a piece of Whatman paper cut to the size of the cassette is probably the simplest way. Never tape anything directly to an intensifying screen.

Development and fixation

Transformation of the invisible latent image into a visible autoradiographic image occurs by the stepwise treatment of the exposed film with a photographic developer, some type of developer stop bath, a photographic fixer solution, and finally a thorough washing of the film. Each type of X-ray film has a complementary set of developer and fixer solutions recommended by the manufacturer. Films can be developed manually in a dark room, in which case the equipment requirements are minimal. Alternatively, films can be developed automatically by machine, if such an apparatus is available. If autoradiographic exposure was carried out at -80° , be sure to warm the film to room temperature before opening the cassette, otherwise the quality of the image and the reproducibility of the assay may be compromised. Films that are developed manually should be washed in running water ($16\text{--}24^{\circ}$) for at least 15 min prior to air-drying to render the highest possible quality autoradiogram.

Protocol: autoradiography

CAUTION: As with all experiments involving isotopes, be sure to follow all departmental guidelines for the proper containment of this material and all safety guidelines to minimize potential health hazards. It is wise to manipulate radioactive filters behind Plexiglas shields.

In order to generate meaningful data, it is critical to standardize both the exposure time and the developing procedure: evaluation of short- and long-term exposures or different developing protocols involving the same blot may suggest very different conclusions.

1. At the conclusion of all posthybridization stringency washes (Chapter 13), place the filter on a piece of Whatman 3MM paper just long enough to blot excessive buffer from the filter.
 2. Wrap the filter in plastic wrap, taking care to avoid wrinkling on the side of the filter to which the probe has hybridized. Fold excess plastic wrap behind the filter so there is no possibility that the X-ray film will get wet.
 3. Hold a Geiger counter probe 10–15 cm above the surface of the plastic-wrapped filter, to estimate the required time for autoradiography. It is not unusual to have very few “clicks” registering just above background when assaying rare mRNAs by Northern analysis. Significantly greater activity may be observed in the region of the filter corresponding to the molecular weight markers if the hybridization cocktail contained an aliquot of probe capable of base-pairing to these sequences, too. High activity distributed over the surface of the filter may be indicative of unacceptable background levels.
 4. Place the plastic-wrapped filter, a sheet of X-ray film, and an intensifying screen together in a cassette, under safelight conditions. Store at -70° overnight.
- Note: Be sure to ascertain that the box containing the unused film is closed before turning on the room lights.*
5. At the conclusion of the exposure period allow the cassette to warm to room temperature. If using an automatic X-ray film developing machine, open the cassette

under safelight conditions and process the film according to the protocol for the developing machine, and then proceed to Step 15. If manually developing the film, follow steps 6–14.

6. Photographic developer⁴ (e.g., Kodak D-19 or D-76) should be prepared in advance and stored in a tightly sealed dark bottle because it is light-sensitive and oxidizes rapidly. Pour enough developer into a photographic developing tray so that the film is submerged when placed into developer. The developer should be clear and colorless when poured from its storage bottle. The developer oxidizes as it ages, acquiring a yellow tint. While not ideal, developer that is *slightly* yellowed will continue to function as long as it is clear, though the quality of the film may be suboptimal. If the developer appears dark yellow or brown, it is completely useless and should be discarded at once.
7. Fill a second photographic tray with some type of stop bath solution, or simply water.
8. Fill a third photographic developing tray with photographic fixer which, like the developer, was prepared in advance and stored in a dark, tightly sealed bottle. The volume of fixer in the third tray should permit the film(s) to be submerged with ease.
9. Take note of the placement of everything in the dark room before turning off the lights. Fumbling around in the dark is counterproductive.
10. Turn the safelight on and room lights off. Wearing gloves, open the X-ray exposure cassette and, before removing the film, bend one corner of the film so that the resulting asymmetry can be used to keep track of the film orientation within the cassette. Do not remove the filter from the cassette or otherwise change its relative position, and be certain that the plastic-wrapped filter is not sticking to the film (occurs frequently).
11. Slide the film into the developer and submerge all at once; do not float the film on the surface of the developer. Begin timing immediately, and gently agitate the developer tray. An exposed film will generally begin to manifest bands within 60 s, though the developing time may be extended to as much as 2.5 min. When the bands begin to appear, note the elapsed time, using it as a guideline for subsequent films. Standardizing developing time gives greater meaning to data generated in different experiments.
12. At the conclusion of the developing period, quickly remove the film from the developer and place in the stop bath. Rinse the film for 20–30 s.
13. Transfer the film to the third tray containing the photographic fixer and submerge. Gently agitate this tray, and notice how, after a few moments, the film itself begins to clear, rendering the characteristic X-ray film appearance. Fixation is complete when the film has *completely* cleared, usually in less than 5 min, though the fixation time should be extended to a full 5–10 min.
14. At the conclusion of the fixation step, it is safe to turn on the room lights. Carefully transfer the developed film into a sink and allow it to rinse in running warm water for at least 10 min. Washing away all of the residual chemicals at this post-developing stage will produce superior quality films. Hang the films to dry.
15. If very little is evident on the film after this initial exposure, then the background is likely to be very good. Set up another exposure, store at -70° for three days, and then repeat the steps for developing the film. If the background on the film is high upon initial exposure, usually indicated by the appearance of an outline of the filter

⁴**Caution:** Always avoid skin contact with photographic developer, stop bath, and fixer, and be certain to wear eye protection when working with these and all chemicals.

paper on the film and with no apparent bands, consider unwrapping the filter paper and repeating the most stringent posthybridization wash⁵; for greater stringency, filters may be washed at 55–60° in 0.1 × SSPE, 0.1% SDS for 10–15 min. Repeat steps pertaining to wrapping the filter and setting up exposure.

16. Follow instructions of the filter manufacturer for the removal of hybridized probe and/or storage of the filter. See Chapter 13 for probe removal suggestions.

Non-isotopic procedures

Non-radioactive methodologies for nucleic acid probe synthesis are remarkably similar to isotopic labeling approaches. The mechanics of hybridization detection, however, are quite different. The non-radioactive methodologies are (1) chemiluminescence, a process in which light is generated by enzymatic modification of suitable substrates; (2) chromogenic (colorimetric) detection, in which an enzymatic activity causes a color precipitate to be deposited directly on the filter at the site of the reaction; (3) fluorescence, in which molecular excitation of certain compounds at one wavelength results in emission of light at a different, characteristic wavelength; and (4) chemifluorescence, in which a fluorogenic substrate is converted enzymatically into a fluorescent product which, upon laser excitation, may be captured using a fluorescence imaging system.

The advent of non-isotopic labeling and detection, first by chromogenic techniques and later by chemiluminescence, was heralded as an alternative to radiolabeled nucleic acid probes that were in widespread use for blot analysis and *in situ* hybridization. As described in Chapter 12, there are many widely available methods for generating non-isotopic probes. Many laboratories have reduced or altogether eliminated use of isotopes in favor of newer methods of labeling and detection. Chemiluminescence in particular has emerged as an excellent method for bypassing the traditional difficulties associated with chromogenic detection, not to mention circumvention of the handling and containment of radioactive materials, as well as the mounting disposal problems associated with their use. In chemiluminescence, light emission is recorded on a piece of X-ray film or via PhosphorImager. The exposure period associated with chemiluminescence is significantly shorter than is necessary for probe detection by autoradiography and is usually on the order of minutes, rather than hours or days. Surprisingly, however, chemiluminescence has yet to achieve widespread acceptance.

Numerous systems that support non-isotopic labeling and detection by chemiluminescence have been developed by many major suppliers of biotechnology reagents and products. In a very general sense, the mechanics of

⁵If using chemiluminescence, a follow-up stringency wash will mandate blocking the filter again, along with all of the subsequent steps for hapten recognition, binding of the enzyme conjugate, and delivery of the appropriate substrate. This constitutes a somewhat labor-intensive process, suggesting that it is best to wash the filter stringently the first time, to minimize the likelihood that additional washes will be required.

non-isotopic detection methodologies involve (1) assembly of an immunohistochemical “molecular sandwich” by the ordered application of a series of reagents, as in the use of biotin (Bio), digoxigenin (DIG), or fluorescein (Fl); or (2) activation of an enzyme that was linked directly to the probe molecules, such as in the Enhanced Chemiluminescence (ECL) system (GE Healthcare).

Biotin

As described in Chapter 12, there are several methods by which probe biotinylation can be achieved. One of the first steps in biotin-based probe detection systems involves the binding of biotin with some form of streptavidin. Very few pairs of molecules exhibit the unusually high affinity that is observed between biotin and streptavidin ($K_d = 10^{-15} \text{ M}^{-1}$). Streptavidin is a tetrameric protein (MW 60,000) isolated from the bacteria *Streptomyces avidinii* and has four biotin binding sites. Unlike avidin (from egg), streptavidin has a neutral isoelectric point at physiological pH, with few charged groups, and contains no carbohydrate. These properties reduce non-specific binding and background problems that would otherwise be experienced and, in so doing, enhance the sensitivity of various assay methods. Many biotin–streptavidin applications pertaining to nucleic acid hybridization and related detection techniques have been described over the years, and it seems that the diversity of applications increases with each passing day (for classic examples, see Leary, *et al.*, 1983; Wilchek and Bayer, 1984; Hofman and Finn, 1985). In the most common detection format, streptavidin is conjugated to either alkaline phosphatase (AP) or horseradish peroxidase (HRP). Subsequently, the choice of substrate for the enzyme will determine the mode of detection. In the case of AP, dioxetane substrates result in the emission of light while BCIP and NBT⁶ together support the formation of a color precipitate. In contrast, HRP also initiates a set of chemical reactions in the presence of H_2O_2 , resulting in the emission of light (luminol-mediated) or the formation of a color precipitate (4-chloro-1-naphthal-mediated).

Digoxigenin

Digoxigenin is widely used as a method for generating non-radioactive probes, and is an alternative to biotinylation. Like biotin, DIG-labeled probes support both chromogenic and chemiluminescence detection formats. Following the posthybridization stringency washes, DIG-labeled probes are detected by enzyme-linked immunoassay, using an antibody conjugate (e.g., anti-DIG-AP conjugate or HRP conjugate). As with systems involving biotin, the emission of light or the formation of precipitate is mediated by enzymatic cleavage of a substrate compatible with the method of detection, either colorimetric or chemiluminescence.

⁶BCIP: 5-bromo-4-chloro-3-indolyl phosphate; NBT, nitroblue tetrazolium.

Fluorescence imaging

As described in Chapter 12, among the more commonly used fluorophores for probe labeling are Cy3 or Cy5, rather than via the older method of fluorescein-labeled nucleotide incorporation. However, this classical method of fluorescence labeling remains in use, and the mechanics of probe detection and signal generation are quite similar to the approach used in conjunction with DIG-labeled probes. Following hybridization, high-affinity antibodies prepared against fluorescein are used to localize the probe, and the conjugation of these antibodies with either AP or HRP will support either chemiluminescence or chromogenic detection.

Regarding direct detection of fluorescein probes hybridized on filter papers: because fluorescein has a relatively high photobleaching rate and the fluorescence is pH-dependent, quantitative measurement of this fluorophore is problematic. While newer instrumentation and newer dyes have overcome some of the intrinsic limitations associated with fluorescein in the filter-based format, detection of fluorescein probes, if used, is best served by enzyme-conjugated anti-fluorescein antibodies.

Direct enzyme labeling

The Enhanced Chemiluminescence (ECL) system (GE Healthcare) features a modified horseradish peroxidase linked directly to probe molecules through a glutaraldehyde linker arm. Thus, the enzyme is present throughout the hybridization period as well as the posthybridization stringency washes. This labeling and detection method precludes the need for any type of hapten recognition, thereby supporting immediate application of chemiluminescence substrates. It remains one of the most rapid methods for labeling and detection and usually produces low-background images. The ECL system is highly respected for Western analysis; the nucleic acid and protein detection formats are in widespread use.

Detection by chemiluminescence

Chemiluminescence is the conversion of chemical energy into the emission of visible light (luminescence) as the result of an oxidation or hydrolysis reaction. This technology provides a very sensitive, cost-effective detection alternative to many radioisotopic and fluorescence techniques, and most chromogenic detection processes (Table 14.2). There are many adaptations of this technology to conventional molecular biology techniques, including Northern analysis, Southern analysis, plaque lifts, and *in situ* colony hybridization. While the underlying principles at work in each system are essentially the same, kits for hybridization detection by chemiluminescence differ mostly in the number of reagents and manipulations required of the investigator in order to generate a signal and, to a lesser extent, in the achievable level of sensitivity.

Labeling and detection of probe molecules by chemiluminescence generally do not influence the parameters that govern nucleic acid hybridization, which are

Table 14.2 Comparison of Chemiluminescence and Autoradiography for Hybridization Detection

Detection parameter	Chemiluminescence	Autoradiography
Sensitivity	Very good to excellent	Excellent
Resolution	Very sharp banding	Doublets often difficult to resolve
Exposure time	Minutes to hours	Days to weeks
Detection	X-ray film or PhosphorImager	X-ray film or PhosphorImager
Potential health hazards	Relatively low	Very high
Probe stability after labeling	Most stable for 1 year	Limited by half-life
Posthybridization handling of filters	Detection buffers required prior to film exposure	No detection buffers required

essentially the same for radiolabeled probes. Further, many of the labeling technologies that support detection by chemiluminescence are identical to those used for chromogenic detection. Differences in the detection systems are related to alternative chemiluminescence substrate formulations. The mechanics for measuring chemiluminescence are similar to those for measuring isotope decay by autoradiography. The relatively intense, albeit brief, emission of light generates a quantifiable image on X-ray film in a fraction of the time needed for detection by autoradiography. Alternatively, phosphorimaging systems can be used for image capture.

Substrates for chemiluminescence

The numerous chemiluminescence detection systems fall into what this Author considers the two schools of chemiluminescence: the alkaline phosphatase school and the peroxidase school⁷. The classical substrates used in alkaline phosphatase systems are 1,2 dioxetane derivatives, including the Lumigen family product line⁸ (Lumigen, Inc., Southfield, MI), CSPD[®], and CDP Star[®] (Applied Biosystems, Foster City, CA), and others. These patented substrate formulations undergo rapid conversion into luminescent form in the presence of calf intestine alkaline phosphatase (Bronstein and McGrath, 1989). The resulting emission of blue light (≈ 470 nm), is in direct proportion to the amount of alkaline phosphatase present (Fig. 14.2).

Traditional formulations containing Lumigen[™] PPD which are compatible with alkaline phosphatase systems include Lumi-Phos[™] 530 (light emission maximum at 530 nm), and Lumi-Phos Plus (light emission maximum at 470 nm). In addition Lumigen APS-5, which is an acridan substrate, offers

⁷ β -galactosidase and luciferase also support chemiluminescence, though their use is most often reserved for reporter gene assays.

⁸Lumigen = 4-methoxy-4-(3-phosphatephenyl)spiro(1,2-dioxetane-3,2'-adamantane), disodium salt.

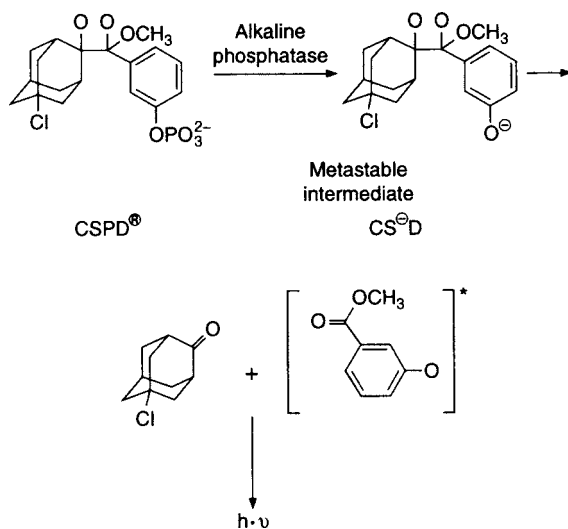


Figure 14.2 Enzymatic dephosphorylation of dioxetane substrate by alkaline phosphatase leads to the metastable phenolate anion which, upon decomposition, emits light at ≈ 480 nm. The light emission increases with time until a constant intensity is attained.

excellent sensitivity down to 10^{-19} moles of alkaline phosphatase and sustained luminescence, making the timing of the detection assay less demanding. Applied Biosystems also produces chemiluminescence enhancers, and these formulations are compatible with many molecular biology applications.

The substrate used in the peroxidase systems is hydrogen peroxide (H_2O_2). By virtue of the chemistry involved, the reduction of H_2O_2 produces oxygen radicals (O_2^-) and H_2 . This, in turn, promotes the luminescent oxidation of the cyclic hydrazide luminol⁹ (Matthews *et al.*, 1985; Thorpe and Kricka, 1986), or any other of a variety of related substrates. The oxidation products are 3-aminophthalate and N_2 , with the associated emission of light (Fig. 14.3). The location and magnitude of hybridization is most often recorded by X-ray film exposure, much as in autoradiography. Other specialized reagents in use for peroxidase-based chemiluminescence detection are Lumigen PS-atto, Lumigen PS-2, and Lumigen PS-3, and the newer Lumigen TMA-3 and Lumigen TMA-6 formulations (Lumigen, Inc.).

Hybridization detection by chemiluminescence often represents the end of a series of steps, which sometimes can be quite elaborate. These steps may include probe labeling, blocking of filters, posthybridization washes, blocking the filters a second time, hapten localization, washing the filter repeatedly, and enzymatic destabilization of suitable substrates. As described previously, the detection mechanics for chemiluminescence involve creating a molecular sandwich. Biotinylated probe molecules, for example, are localized by reaction with a streptavidin that is conjugated to either AP or HRP. Alternatively, DIG-labeled

⁹Luminol = 5-Amino-2,3-dihydro-1,4-phthalazinedione.

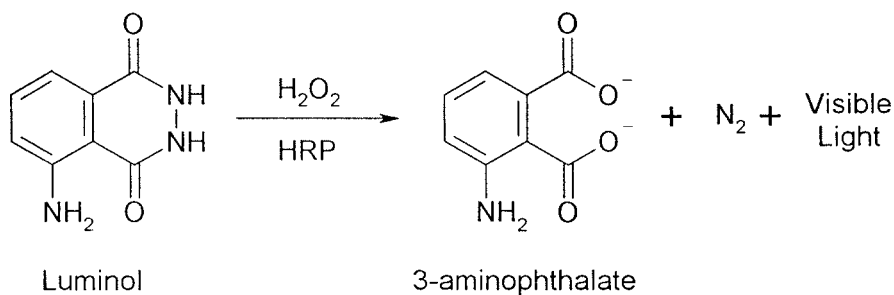


Figure 14.3 Horseradish peroxidase (HRP), in the presence of H_2O_2 and luminol results in the production of visible light. The emission of light peaks within in few minutes after application of the detection reagents and then falls precipitously.

probes are localized by anti-DIG antibodies, which exist as an enzyme conjugate. The same is true of the anti-fluorescein antibodies used to detected fluorescein-labeled nucleic acid probes. Then, upon incubation with the appropriate substrate, enzyme activation results in the emission of visible light. To generate a record of the intensity and location of light emission, substrate-soaked filter papers are quickly plastic wrapped and exposed to X-ray film using the same types of exposure cassettes describe in the “Autoradiography” section earlier in this chapter. As there is no isotope involved, the resulting image might best be thought of as a chemilumigram, rather than an autoradiogram, though most people simply refer it as “the film”.

The most common difficulties with chemiluminescence as a detection technology are the incidence of unacceptably high levels of background, or no signal at all. In the case of the former, this is most often due to (1) incomplete washing of filters posthybridization and during the detection steps; (2) incomplete blocking of the filter prior to addition of the enzyme conjugate; (3) allowing the filter to dry out after the detection procedure has been initiated; and (4) the use of powdered latex gloves (problematic during processing of X-ray films, too). In this lab, stringent posthybridization washes, strict attention to the manufacturer’s recommended detection protocol, and the use of powder-free gloves have produced remarkably good quality exposures, with reproducible high levels of sensitivity. In the case of the latter, failure to detect a signal (including the positive control) is most commonly due to confusion and unfamiliarity with the unique mechanics of chemiluminescence, compared to autoradiography.

Chromogenic detection procedures

Many of the chromogenic detection systems that appeared in the mid-1980s failed to gain widespread acceptance for one reason or another. Chromogenic systems lost popularity with the debut of many refined systems for solid-phase (nucleic acids immobilized on a filter) detection by chemiluminescence. One noteworthy exception is the DIG Labeling System (Roche Applied Science)

which has become well respected as a method for non-isotopic *in situ* hybridization detection.

Chromogenic detection and procedures for chemiluminescence are often similar up to the addition of the very last reagent, the enzyme substrate. In the final step of a typical AP-based chromogenic detection, the enzyme-mediated dephosphorylation of BCIP and subsequent interaction with NBT result in the formation of a blue/purple precipitate, which is deposited directly on the filter membrane. This reaction occurs only at sites that correspond to hybridization events. There is no film exposure involved as there is no autoradiography or chemiluminescence. The chromogenic detection process may often require up to several hours to reach completion. The low sensitivity observed in some chromogenic systems is due to the fact that the activity of alkaline phosphatase is inhibited as the color precipitate forms and is deposited on the filter. Enzyme inhibition in this manner quickly compromises the quantitiveness of the assay, resulting in unreliable, non-representative data. The final step of an HRP-mediated, chromogenic detection involves the application of H_2O_2 and either 4-chloro-1-naphthol or TMB¹⁰.

A bit of an anecdote: On one occasion in this lab, plaque lift filters were being processed with the intention of performing chromogenic detection, though a decision made at the last minute resulted in the substitution of chemiluminescence reagents (hydrogen peroxide and luminol) rather than hydrogen peroxide and 4-chloro-1-naphthol. This last minute procedural change had no negative impact on the outcome of the screening and allowed the successful identification of salient clones in the library (Fig. 14.4). While not a recommended standard procedure, this information is included here to underscore the similarity of these methods.

In order to reprobe filters that were previously subjected to chromogenic detection methods, very harsh washes of the filter are required, to remove the precipitate (and probe) from the filter. Each type of filter membrane and non-isotopic detection system has a set of recommended guidelines for stripping previously hybridized filters, and preparations for reprobing. Typical procedures are presented in Chapter 13.

Digital imaging systems

Now standard fare in most molecular biology laboratories, digital image analysis of data at some level has added a new dimension to both the sensitivity that is achievable and the accuracy of interpretation. While digital imaging systems often represent a significant investment, the purchase of such devices is a one-time expense, upgrades and service contracts notwithstanding. If this type of instrumentation is beyond the means of a laboratory or institution, one can add a contemporary electronic twist to the analysis of one's data by digitizing X-ray films or photographs of gels and using stand-alone image analysis software,

¹⁰TMB = 3,3',5,5'-tetramethylbenzidine.

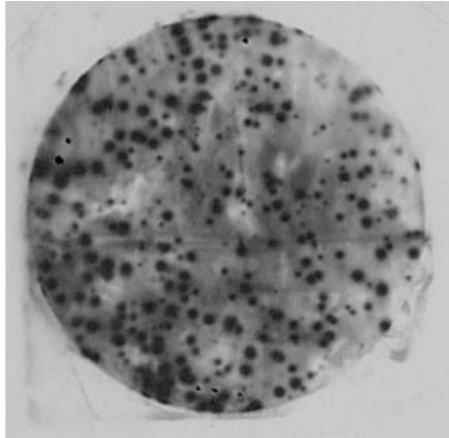


Figure 14.4 Plaque screening assay by chemiluminescence. Luminol was used as a horseradish peroxidase substrate to recover a clone from a cDNA library. This image represents a five-minute exposure. As expected, nearly every plaque shows up as positive because this was the third-round of clone replating and clone enrichment after the initial library screening. All of the clones represented on the filter are now genetically identical, any one of which could then be retrieved from the corresponding plate for sequencing.

such as Gel-Pro (Media Cybernetics) to make unbiased measurements of the molecular weight and mass of observed bands. Many vendors now offer digital imaging systems that are able to capture images from gels and blots for completely film-less electronic analysis and archiving. Some of these systems are mentioned in Chapter 10 and a few of salient points are reiterated here with emphasis given to blot analysis and the elimination of X-ray film altogether.

A PhosphorImager is a sophisticated, very sensitive imaging device, the primary function of which is to produce digital images of the gels and blots. The first model became available in the late 1980s from a company then known as Molecular Dynamics and was originally designed for the quantification of radioactivity in gels, including ^{32}P , ^{33}P , ^{35}S , ^{14}C , ^{125}I , and ^{131}I . Today, two of the widely known and respected PhosphorImagers are the TyphoonTM and Storm[®] models (GE Healthcare). Depending on the model, PhosphorImagers are capable of quantification of fluorescence (e.g., ethidium bromide, SYBR Green, SYBR Gold), chemifluorescence, and radioactive decay. Some systems are capable of imaging gels stained with Coomassie blue and silver stained gels, too. Analysis has also been extended to the realm of microarrays.

There are numerous advantages associated with PhosphorImagers, and the mode of image capture is unique. First and foremost, there is no X-ray film involved. In the case of blot analysis, images are generated on special screens that can be used over and over and, with proper care, will last indefinitely. The exposure time can be as little as one-tenth the time required for X-ray films. The PhosphorImager detection technology is completely compatible with the qualitative and quantitative assay of nucleic acids, proteins, and any other

macromolecule that carries a fluorescent or radioactive tag. Best of all, most investigators agree that, by comparison with X-ray film, the sensitivity associated with PhosphorImaging is at least 10 times greater. Moreover, the elimination of X-ray film also means the elimination of the very narrow linear response that is an intrinsic shortcoming of all photographic films. Typically, PhosphorImagers offer a linear response of five logs, resulting in very quantitative data that highlights the true variations among samples.

Because of the popularity of PhosphorImagers and the fact that a department or institution is likely to have only one such instrument available because of the cost, it is very common for users to sign up well in advance for a time to use the machine. In high-throughput laboratories and/or laboratories with several graduate students, the time slots available to use the machine may not be until the wee hours of the morning!

References

- Bonner, W.M. (1987). Autoradiograms: ^{35}S and ^{32}P . *Methods Enzymol.* **152**, 55–61.
- Bonner, W.M. and R.A. Laskey. (1974). A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. *Eur. J. Biochem.* **46**, 83–88.
- Bronstein, J. and P. McGrath. (1989). Chemiluminescence lights up. *Nature* **338**, 599–600.
- Hofman, K. and F.M. Finn. (1985). Receptor affinity chromatography based on the avidin–biotin interaction. *Annals N.Y. Acad. Sci.* **447**, 359–372.
- Laskey, R.A. and A.D. Mills. (1975). Quantitative film detection of ^3H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* **56**, 335–341.
- Laskey, R.A. and A.D. Mills. (1977). Enhanced autoradiographic detection of ^{32}P and ^{125}I using intensifying screens and hypersensitized film. *FEBS Lett.* **82**, 314–316.
- Leary, J.J., D.J. Brigati and D.C. Ward. (1983). Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose. *Proc. Natl. Acad. Sci.* **80**, 4045–4049.
- Matthews, J.A., A. Batki, C. Hynds and L.J. Kricka. (1985). Enhanced chemiluminescent method for the detection of DNA dot-hybridization assays. *Analyt. Biochem.* **151**, 205–209.
- Randerath, K. (1970). An evaluation of film detection methods for weak β -emitters, particularly tritium. *Analyt. Biochem.* **34**, 188–205.
- Swanstrom, R. and P.R. Shank. (1978). X-ray intensifying screens greatly enhance the detection by autoradiography of the radioactive isotopes ^{32}P and ^{125}I . *Analyt. Biochem.* **86**, 184–192.
- Thorpe, G.H. and L.J. Kricka. (1986). Enhanced chemiluminescent reactions catalyzed by horseradish peroxidase. *Methods Enzymol.* **133**, 331–354.
- Wilchek, M. and E.A. Bayer. (1984). The avidin–biotin complex in immunology. *Immunology Today* **5**(2), 39–43.
- Zagursky, R.J., P.S. Conway and M.A. Kashdan. (1991). Use of ^{33}P for Sanger DNA sequencing. *BioTechniques* **11**, 36–38.

15 Quantification of Specific mRNAs by Nuclease Protection

Rationale	321	Protocol: transcript quantification	
Basic approach	322	by RNase protection	335
Probe selection	327	Troubleshooting	337
Optimization suggestions	329	References	338
Potential difficulties	330		
Protocol: transcript quantification			
by S1 analysis	332		

Rationale

The study of rare or low abundance transcripts is a frequently reiterated theme in the molecular biology laboratory. The various permutations of PCR, without a doubt, offer the most compelling quantitative expression profile of a sample, though PCR is not without limitations of its own. A solidly designed investigation should be formatted in such a way that data generated by PCR can be validated using a non-PCR based technique.

One approach that can be used to generate or validate quantitative data is generically known as a nuclease protection assay. The crux of the assay is the formation of double-stranded hybrids between probe and target RNA by solution hybridization under highly stringent conditions, rather than using filter-immobilized RNA, after which non-hybridized and non-specifically-hybridized molecules are eliminated by nuclease digestion. This assay was developed years before the advent of PCR to alleviate the lack of quantitiveness associated with Northern analysis, caused by the inaccessibility of some of the membrane-bound RNA to probe sequences. Whenever a filter membrane is involved, complete hybridization is difficult, if not impossible. Further, the sensitivity and resolution of Northern analysis are often compromised by the inability to load large amounts of RNA on a gel, by the inefficient transfer (blotting) from gel to filter membrane, and by non-specific hybridization between probe and transcripts exhibiting limited homology.

Transcript assay by nuclease protection is a very sensitive and precise method for quantifying specific transcripts. It is widely accepted as a reliable, independent method that can confirm results generated by RT-PCR, such as identification of a putative transcription start site, changes in transcript abundance, and related techniques. The investigator has the choice of using total cellular, total cytoplasmic, poly(A)⁺ mRNA, or any other RNA subset as starting target material, though the enhanced sensitivity of this approach usually does not mandate poly(A)⁺ selection. Hybridization is followed immediately by exposure to a single-strand-specific nuclease that ordinarily has no affinity for double-stranded molecules, i.e., either DNA:RNA or RNA:RNA hybrids. Thus, any probe sequences or target RNA molecules that do not participate in duplex formation are digested, and the resultant protected fragments recovered by ethanol precipitation. The size and mass of these hybrid molecules are then deduced by denaturing polyacrylamide gel electrophoresis, and the signal intensity is quantified by in-gel autoradiography (Figs. 15.1 and 15.2) or PhosphorImaging. In one variant of this technique, the abundance of the undigested molecules is assessed by liquid scintillation counting instead. The only short-coming of nuclease protection is the fact that the protected fragments are not indicative of the size of the native transcript. This information can only be ascertained by Northern analysis or sequencing of a full-length cDNA clone.

By virtue of the solution hybridization of probe and target, followed by digestion of all other sequences that do not participate in duplex formation, the signal-to-noise ratio is boosted, often dramatically. The investigator can expect at least a 10-fold increase in sensitivity over conventional blot analysis and the ability to reproducibly detect as little as 5 femtograms (fg) of target RNA¹ after a 3-day X-ray film exposure using high specific activity probes. Moreover, with the incorporation of a 5' label via the kinasing reaction, probes of high specific activity (>10⁹ cpm/μg) can be generated for the precise mapping of the 5' end of the transcript (Boorstein and Craig, 1989). This method has also been used in the past for mapping of exon-intron boundaries (Weaver and Weissman, 1979). Any transcript that can be studied quantitatively by standard Northern analysis is easily detectable by a more sensitive nuclease protection assay. This is especially convenient when working with small amounts of RNA from several different samples.

Basic approach

A variety of nucleases and approaches have been used to quantify and characterize specific transcripts in greater detail than can be derived from Northern analysis. The selection of alternative approaches is expressly contingent upon the

¹Although autoradiography is considered to be the preferred method for maximum sensitivity, successful non-isotopic variations of both assays have been reported (Rosenau et al. 2002; Kris et al., 2007).

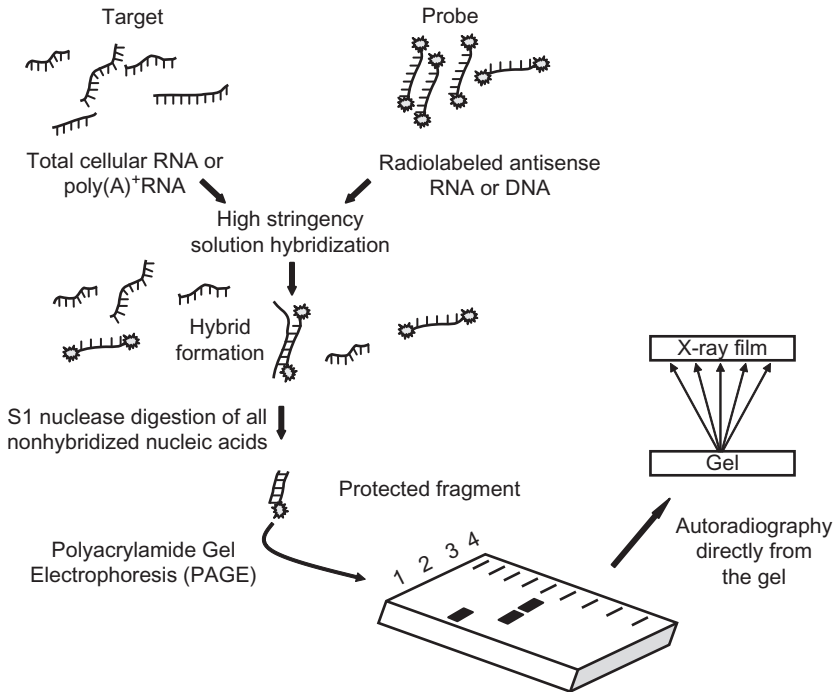


Figure 15.1 The S1 nuclease assay for the quantification of specific RNA species. Purified RNA is hybridized in solution with a labeled probe sequence to form thermodynamically stable hybrid molecules. Any RNA or probe molecules that do not participate in the formation of hybrid molecules are digested away by the single-strand-specific nuclease S1, followed by electrophoresis of the intact hybrid molecules. The size and abundance of protected RNAs are then deduced by autoradiography, performed directly from the gel. Lane 1: undigested probe. Lanes 2 and 3: experimental samples. Lane 4: molecular weight standards.

questions being asked in a particular investigation. These enzymes include nuclease S1 (Ando, 1966; Vogt, 1973), mung bean nuclease (Kowalski *et al.*, 1976), and nuclease VII (Chase and Richardson, 1974a,b), and a combination of RNase A and RNase T1 (Zinn *et al.*, 1983; Melton *et al.*, 1984; Lee and Costlow, 1987). The associated techniques include primer extension (McKnight and Kingsbury, 1982; Jones *et al.*, 1985; Kingston, 1988), the nuclear runoff assay (Marzluff, 1978; Greenberg and Ziff, 1984), the ribonuclease protection assay (Zinn *et al.*, 1983) and the S1 nuclease protection assay (Berk and Sharp, 1977; 1978). To a great extent, PCR has replaced some of these approaches, though the S1 protection assay and RNase protection assay (RPA) remain popular and reliable techniques. The key advantage that these techniques offer is that they do not require a reverse transcriptase step (see Chapters 18 and 19 for details).

The basic methodology of nuclease protection accommodates denatured double-stranded DNA probes as well as RNA probes. Hybridization is usually

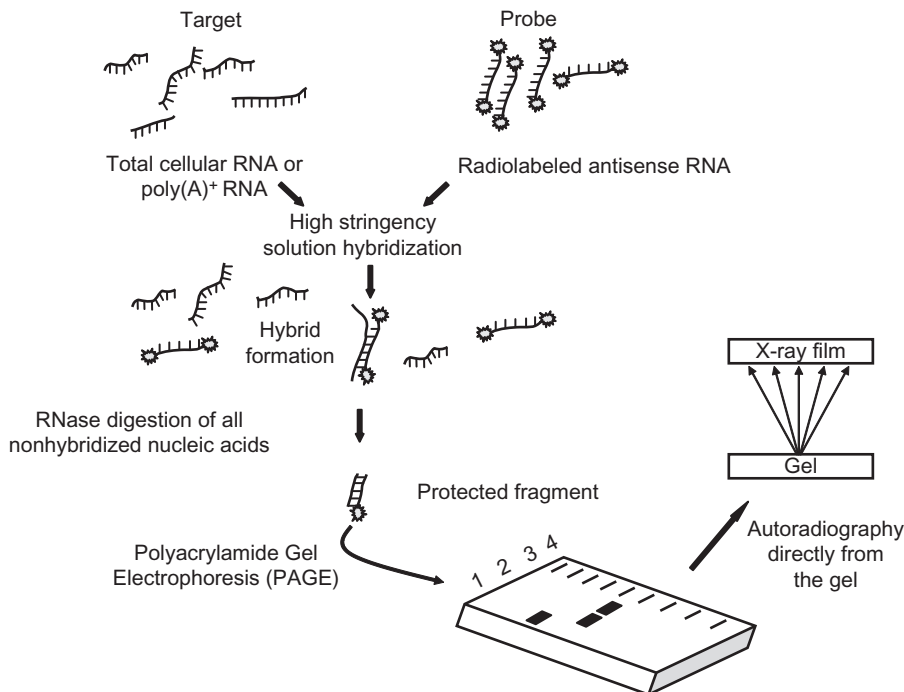


Figure 15.2 RNase protection assay for the quantification of specific RNA species. Purified RNA is hybridized in solution with a labeled antisense probe sequence to form thermodynamically stable double-stranded RNA molecules. Any RNA or probe molecules that remain single stranded are digested by an RNase cocktail. Following electrophoresis, the size and abundance of protected RNAs are then deduced by autoradiography, performed directly from the gel. Lane 1: undigested probe. Lanes 2 and 3: experimental samples. Lane 4: molecular weight standards. The general approach is identical to that for the S1 nuclease assay.

performed in buffer containing high concentrations of formamide, which maintains a high level of stringency. Keeping the hybridization volume small and the concentration of nucleic acid high allows the hybridization to go to completion in a few short hours. The formamide-based environment tends to produce very low background by promoting the formation of RNA:DNA duplexes rather than the renaturation of complementary strands of DNA. The use of single-stranded probes, therefore, is clearly preferred, because of the potential for double-stranded probe renaturation issues. In addition, the hybridization temperature must be slightly below the melting temperature (T_m) of the hybrid molecules that are expected to form. Because of the posthybridization removal of single-stranded regions (mismatches and overhangs) in DNA:DNA, RNA:DNA, and RNA:RNA structures, failure to identify this “window of hybridization” will very likely result in suboptimal probe:target duplex formation and thus yield misleading data.

Nuclease S1 is a naturally occurring glycoprotein isolated from the filamentous fungus *Aspergillus oryzae*. It is the least expensive nuclease for this type of analysis and was once used interchangeably with mung bean nuclease in RNA mapping assays. The S1 nuclease is a zinc-dependent² enzyme (Ando, 1966; Vogt, 1973) that has a high specificity for single-stranded or imperfectly base-paired polynucleotide structures. Any single-stranded DNA or RNA regions are hydrolyzed exo- and endonucleolytically, yielding 5'-mononucleotides and 5'-oligonucleotides and the hydrolysis of DNA is about five times faster than is observed for RNA. Optimal nuclease activity is observed around pH 4.5, under conditions of relatively high ionic strength³ and it is essentially inactive at pH 7.2 (Linn and Roberts, 1982). S1 nuclease exhibits thermostable qualities (Ando, 1966), although in many applications, incubations are conducted at 37°. Moreover, the enzyme is resistant to denaturants (Vogt, 1973; Hofstetter *et al.*, 1976), including formamide, sodium dodecyl sulfate (SDS), and urea, thereby enhancing its usefulness.

While there are many advantages of an enzyme possessing these characteristics, there are potential difficulties associated with the use of the S1 nuclease enzyme as well. S1 nuclease is an aggressive enzyme⁴ that is often difficult to control. For this reason alone many investigators prefer the RPA over S1 protection. This part of the protocol, more than any other, is the step that is most likely to require optimization, with a typical range of 100U to 200U per ml of reaction volume. Excessive amounts of S1 nuclease activity will result in the destabilization and rapid digestion of hybrid molecules and naturally unstable AT- and AU-rich regions are particularly susceptible to S1 attack. Nicking, if it does occur, will become evident by the presence of an unexpected lower molecular weight band(s), a situation that can be remedied by reducing the amount of enzyme used for the digest. Further, because of the low pH requirement for optimal activity of this enzyme, acid depurination of double-stranded DNA is an expected consequence of prolonged incubation in such an environment.

The basic approach in the RPA (Fig. 15.2) is precisely the same as the S1 assay (Fig. 15.1): a probe that is complementary to RNA transcripts under investigation is used in solution hybridization, followed by digestion of excess probe and all transcripts that are not locked up in a protected fragment. From a mechanical point of view, one should note that in the S1 assay, either a DNA or an RNA probe can be used, because S1 nuclease is single-strand specific. Table 15.1 presents a comparison of the closely related S1 nuclease- and

²CO²⁺ is also effective as a cofactor (Vogt, 1973).

³At high concentrations of enzyme, as required in some applications, a concomitant increase in the ionic strength of the assay environment is in order, to minimize the nicking of double-stranded molecules that is often observed at higher enzyme concentrations. Nicking of double-stranded DNA is suppressed at 200mM NaCl (Vogt, 1973; Vogt, 1980) and at a slightly higher than optimal pH (Weigand *et al.*, 1975).

⁴The aggressive nature of the S1 nuclease enzyme was partially responsible for poor overall conversion of mRNA into double-stranded cDNA and failure to make full length cDNA in the old days of molecular biology, i.e., before the popular method of Gubler and Hoffman (1983) vastly improved the efficiency of cDNA synthesis. For the young scientists reading this, that is how we used to make cDNA in the days before PCR.

Table 15.1 Comparison of S1 and RNase Protection Assays

	S1 nuclease assay	RNase protection assay
Target	RNA	RNA
Probe	DNA or RNA	RNA
Nuclease	S1 nuclease	RNase A + RNase T1
Background potential	Low to moderate	Moderate to high
Ability to control	Very difficult	Moderately difficult
Average limit of sensitivity	100 fg	50 fg

Table 15.2 Comparison of RNase A and RNase T1

	Typical stock solution	Substrate ¹	Action	Natural source
RNase A	2 mg/ml	Single-stranded RNA	Targets pyrimidines (C and U); cleaves phosphodiester bond 3' to Py/pN; generates 3' phosphate group.	Bovine pancreas
RNase T1	10U/ml	Single-stranded RNA	Endonuclease that targets guanine nucleotides; attacks the phosphodiester bond 3' to GpN; generates 3' phosphate group.	<i>Aspergillus oryzae</i>

¹Attack of double-stranded RNA occurs when exposed to excessive quantities of these enzymes for extended periods.

RNase protection techniques. In the RPA, however, digestion is performed with a combination of RNase A and RNase T1, mandating the use of an anti-sense RNA probe. RNase A is an endoribonuclease that cleaves single-stranded RNA, specifically on the 3' side of pyrimidine (cytosine and uracil) nucleotides (Py/pN). The resulting degradation products are all classified as 3' pyrimidine phosphates, meaning that each molecule ends in a pyrimidine with a 3' phosphate group. RNase T1, also a single-strand-specific endoribonuclease, cleaves the phosphodiester bond at the 3' end of guanine nucleotides (GpN) and, like RNase A, produces oligonucleotides with terminal 3' phosphates. The actions of these enzymes are summarized in Table 15.2.

From a strategy perspective, it is rather stylish to multiplex nuclease protection assay reactions, that is, to use more than one probe in a single reaction tube. The notion is that as long as the probes manifest similar thermodynamic behavior and the sizes of the protected fragments are different, they can be assayed simultaneously, generating a different size band for each protected fragment, all of which appear in a single lane on a gel (Fig. 15.3).

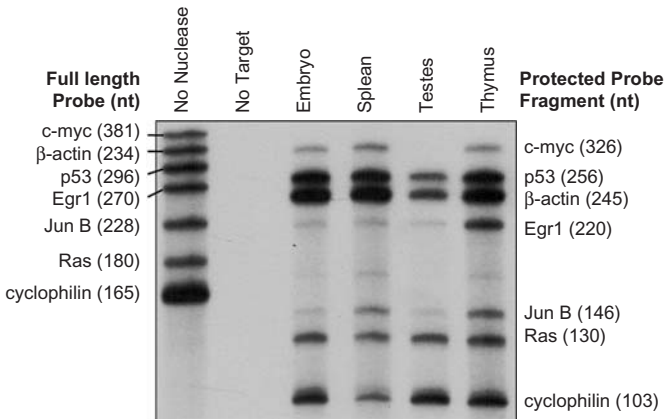


Figure 15.3 Ten micrograms of various mouse tissue total RNAs were hybridized overnight with approximately 50,000 cpm each of seven distinct probe transcripts. Nuclease digestion, product separation on a denaturing 6% acrylamide gel, and a 4 h exposure to film at -80° were used to assess hybridization levels. The radiolabeled probes were synthesized in 5 μ l transcription reactions using Ambion's MAXIScript Kit with $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ (800 Ci/mmol, 10 mCi/ml) and gel purified prior to hybridization. The specific activities of the cyclophilin and β -actin probes were reduced 20- and 200-fold, respectively, by adding appropriate amounts of non-radioactive UTP to the transcription reaction. Courtesy of Ambion.

Probe selection

Many of the potential problems associated with spurious hybridization and probe renaturation may be avoided by the use of single-stranded nucleic acid probes. With the ready availability of highly efficient, cost-effective, complete systems for *in vitro* transcription, large quantities of single-stranded probes of well-characterized sequence and high specific activity can be produced in short order. This is the favored method for generating single-stranded probes for nuclease protection analysis. Alternative, though less frequently used, methods for single-strand probe synthesis for this application include oligonucleotide synthesis, and the nearly antiquated methods of cloning into M13 and asymmetric PCR. Thus, when single-stranded probes are generated, the potential for probe renaturation becomes a non-issue, although intramolecular base-pairing can still occur under conditions of relaxed stringency.

Single-stranded probes generated by *in vitro* transcription almost always contain at least a small quantity of the DNA template from whence the probe was derived. When present, both the DNA template and contiguous vector sequences can (1) potentially hybridize to full-length probe sequences, protecting them from nuclease degradation; (2) promote artifactual hybridization; and (3) reduce the concentration of available probe. High stringency hybridization alleviates these phenomena to some extent, though ideally the transcription vector should be removed by DNase I digestion, followed by probe precipitation

to remove the degraded template. Last, failure to maintain a (ten-fold) molar excess of probe over target RNA will certainly put the sensitivity of the assay in jeopardy.

Another major difference between the Northern analysis and the nuclease protection assays involves interpretation of the banding distribution observed after electrophoresis and autoradiography (Fig. 15.4). In the classical Northern analysis, the observed signal is interpreted as corresponding to the native size

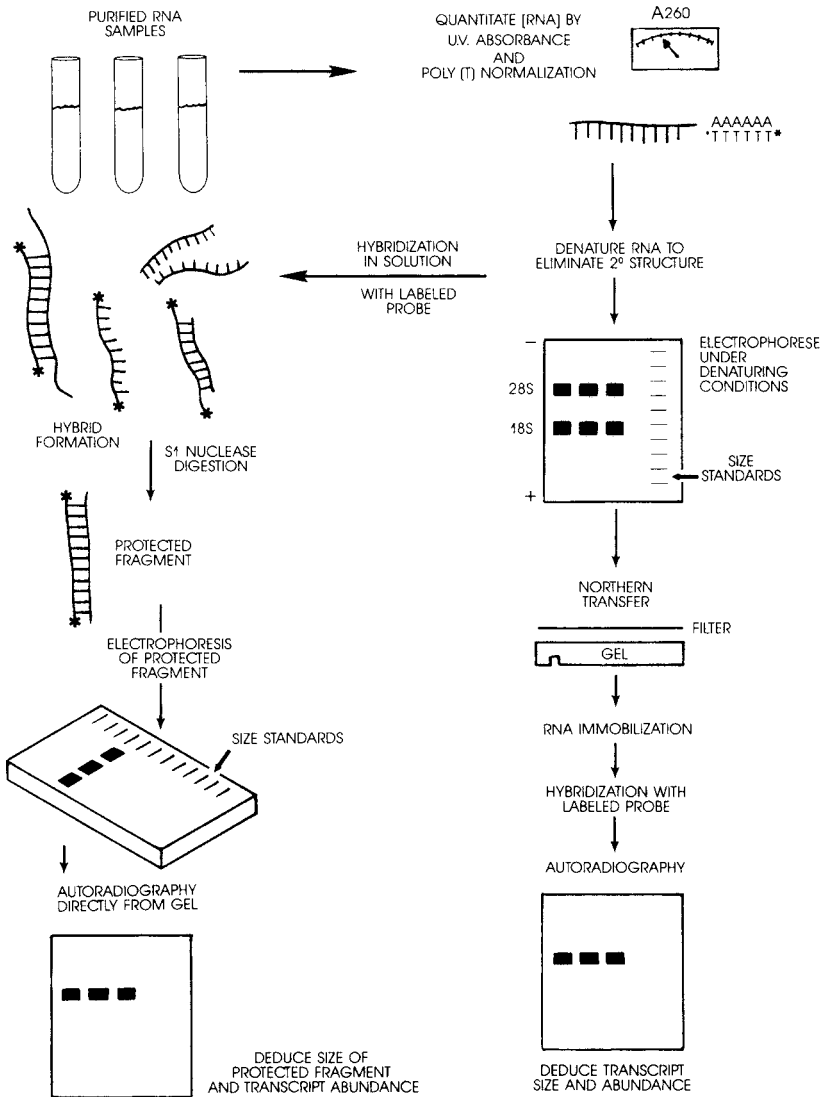


Figure 15.4 Quantification of specific RNA transcripts by nuclease protection compared with conventional Northern analysis.

of the transcript. For example, in humans, the major mRNA produced by transcription of the *c-myc* oncogene is 2.4 kb. As such, this transcript is expected to lie between the very abundant ribosomal RNAs (rRNAs), just above the 18S species. In contrast, the resulting hybrid molecule generated by nuclease protection analysis can only be as long as the probe, and is usually shorter than that because nucleotides from both the 5' and 3' ends of probe and target molecules that do not directly participate in duplex formation are digested. The mechanics of the nuclease protection assays make these methods tolerant of partially degraded RNA samples as long as the integrity of the section of the target molecule recognized by the probe has not been compromised; recall that a single break in a transcript results in that molecule not being detected by standard Northern analysis. In a well-planned assay, the length of the hybrid (the protected fragment) after nuclease digestion should be detectably shorter than non-hybridized or renatured (probe with template) molecules. Under standard assay conditions, electrophoresis of nuclease-treated samples can easily resolve molecules differing in size by only a few nucleotides.

The manner in which probes are labeled is an important concern, especially when one is attempting to characterize a transcript for the first time. End-labeled probes (5' or 3') are *not* preferred and should be avoided, especially in the initial characterization of probe. If the nucleotide containing the (end)-label does not directly participate in duplex formation, the label will be digested away, rendering the hybrid undetectable. This concern can be quickly alleviated with the use of continuously labeled probes (probes that are labeled along the length of the molecule) and probes that are 150 to 450 bases tend to work best.

As indicated previously, single-stranded probes can be easily generated by *in vitro* transcription. Keep in mind that RNA probes must be handled using excellent RNase-free technique, and that the extremely high specific activity associated with these (and other) probes may result in radiolysis⁵ of the probe itself. Fragmented probes translate into increased background and non-specific hybridization. The investigator is well advised to monitor the integrity of the probe (labeled or not) on a gel prior to each use.

Optimization suggestions

Once the type of probe and method of labeling have been established, empirical determination of the hybridization parameters will support the most quantitative nuclease protection assay with the highest fidelity. The following suggestions may facilitate nuclease protection analysis with the highest reproducible levels of sensitivity and resolution. Because nuclease protection reagents are readily available in kit form, some of the optimization procedures listed here may not

⁵Radiolysis is the phenomenon by which radiolabeled molecules are actually degraded by decay of the label they carry. This occurs most often with probes that have been labeled to extremely high activity ($>1 \times 10^9$ cpm/ μ g) and are then stored for more than a few days prior to use.

be necessary. If needed, an optimization regimen is usually required only once. Tweaking the reaction, however, is almost always required for each new probe.

1. The stable formation of target:probe hybrids must be favored. In the absence of detailed information about the behavior of a particular probe in a nuclease protection assay, try hybridization first at 50°. Then, repeat the assay at various temperatures ($\pm 5^\circ$, $\pm 10^\circ$, and so forth) to determine the optimum stringency level for each probe. The temperature should be high enough to allow hybrid formation but not so high as to cause the dissociation of duplex molecules – try adjusting the hybridization temperature to 5° below the T_m .
2. With regard to probe stability, one should beware of AU-rich domains. The intrinsic instability of these regions can result in incomplete hybridization, thereby contributing to a drastic reduction in signal. To an extent, this difficulty can be alleviated by reducing the formamide concentration in the hybridization buffer, though a loss of probe specificity may result. Following hybridization, an incubation at 15° for 30 minutes in diluted hybridization cocktail *prior* to nuclease degradation permits more complete annealing of AU-rich regions, without favoring non-specific hybridization that might otherwise accompany an increase in reaction volume under these conditions (Ausserlechner *et al.*, 1998). Following this incubation, nuclease digestion proceeds as usual.
3. Hybridization kinetics must be driven to completion, and keeping the nucleic acid components of the reaction as concentrated as possible at all times will help. This can be accomplished by the inclusion of carrier RNA (e.g., junk yeast RNA or tRNA). To determine this empirically, remove aliquots at specific intervals and digest each with S1 nuclease as described below. When the hybridization signal intensity no longer increases, the hybridization reaction has gone to completion.
4. Another rather common problem associated with transcript quantification by nuclease protection pertains to the use of antisense transcripts as probes which often consists of a heterogeneous mixed bag of non-uniform-length probe molecules. This, in turn, generates smaller than expected protected hybrids, manifested as more than one band upon autoradiography. This symptom can be alleviated to an extent by gel purification of the newly transcribed cRNA (antisense RNA) prior to use as a probe. Further, it is wise to electrophorese an aliquot of undigested probe alone, in a lane adjacent to the experimental samples, to convince data critics that the assay is functioning properly.

Potential difficulties

Difficulties associated with nuclease protection assays generally fall into one or more of the following categories:

1. Excessive background. This is more of a problem when working with a new probe or set of reagents. Excessive background appears as long smears in the lanes of the gel, usually due to incomplete nuclease digestion. It is normal, however, to observe fairly broad smears at the leading edge of the gel far beyond the location of the protected fragments. One common mistake that can result in elevated background levels is waiting too long to place the samples at the hybridization temperature after the high temperature RNA denaturation step. This interval affords the opportunity

for RNA molecules to regain some degree of secondary configuration, or to base-pair to one another or to non-specific target sequences inappropriately. Efficient handling of samples is essential in this step. Some investigators prefer to program a thermal cycler to move the samples between the denaturation and hybridization temperatures rather than involving two (or more) water baths.

Alternatively, the use experimental RNA that is partially degraded is an invitation to the probe to be less discriminating. Occasionally, “noise” will be observed, resulting from the breakage (due to radiolysis or nuclease degradation) of older probes. Infrequently, smearing in the lanes of the gel may be caused by adding a huge excess, as opposed to a “healthy” excess, of probe to the hybridization reaction mixture. More often than not, the additional radioactivity will be reflected as a large smear at the leading edge of the gel (in the form of free nucleotides), assuming that the enzyme actually does degrade the non-hybridized probe.

Last, the investigator may wish to check the activity of the S1 nuclease or RNase cocktails to ensure that its use is consistent with the recommendations of the manufacturer. Check the expiration date, too. Be certain also that the integrity of the experimental RNA has not been compromised (see Chapter 6).

2. No signal. There are six common reasons for lack of signal at the end of a nuclease protection assay:
 - a. The probe was not labeled to sufficiently high specific activity. Further, in the case of end-labeling, the nucleotide(s) containing the label may have been removed from the hybrid upon nuclease digestion. Although continuously labeled probes (Chapter 12) are less likely to experience this phenomenon, it remains a concern unless it is known with absolute certainty that the actual labeled nucleotide(s) participate in duplex formation.
 - b. The experimental RNA may have been degraded. While it is true that nuclease protection assays are more tolerant of *partially* degraded RNA samples, the degradation cannot extend into the region of the RNA which is complementary to the probe. Always assess the integrity of experimental RNA before doing anything with it.
 - c. Hybridization stringency may have been too high for the type of probe used. The preferred way to remedy this is to lower the hybridization temperature drastically in order to determine if the components of the assay will permit hybridization at all. Then, gradually increase the temperature until the optimal conditions have been determined.
 - d. The precipitated pellet (sample RNA, probe, carrier RNA) may not have been thoroughly resuspended and, therefore, some or most of the sample was unable to participate in the hybridization reaction. Keeping the pellet *slightly* damp with ethanol will assist in dispersing and dissolving it in the hybridization buffer at the onset of the assay. In extreme cases, it is also possible that the nucleic acid pellet may have been lost during the postprecipitation ethanol washes.
 - e. The specific mRNA of interest may be below the level of detection in the cells under investigation. While selection of poly(A)⁺ mRNA prior to S1 analysis may be marginally helpful, it is also potentially counterproductive: the loss of poly(A)⁺ material because of the intrinsic inefficiency of the selection process may further under-represent low abundance mRNAs of interest, and will exclude poly(A)⁻ species. Simply adding a greater starting mass of total RNA to the assay tube may help. Alternatively, the investigator may be aware of a method for inducing or superinducing the specific gene(s) of interest. The hybridization

competence of an RNA sample can be tested using a different probe, ideally a probe for a fairly abundant housekeeping gene, to demonstrate that the sample was at least capable of hybridization to something (anything).

- f. Inadvertent use of a “sense”, rather than an “antisense”, probe.
3. Background is excellent but there are too many bands. If this is the only difficulty with this assay, only minimal adjustments should be needed to optimize it. The appearance of multiple, well-defined bands can result from the presence of DNA that is homologous to the probe. cRNA molecules produced by *in vitro* transcription are often contaminated with plasmid template sequences. A vector or template sequence generally bands closer to the origin (the wells of the gel) because of its much greater size and the fact that it was linearized in order to support the *in vitro* transcription reaction. Clearly, it is important to destroy the template by DNase I digestion. Alternatively, relaxed hybridization stringency may have tolerated probe hybridization to related transcripts, from the members of the same gene family. Depending on the degree of homology and distribution of mismatches, fairly interesting (or frustrating!) banding patterns may be observed. To explore this possibility, conduct an identical hybridization 5° above the hybridization temperature that resulted in the appearance of multiple bands. This simple procedural change is likely to destabilize mismatched hybrids and reduce or eliminate the phantom bands. Lastly, the appearance of multiple bands may also occur in the S1 nuclease assay when denatured, double-stranded DNA is used as a probe due to probe renaturation. For this reason, many investigators simply prefer to use RNA probes directed against RNA target sequences and the RNase protection assay.

Protocol: transcript quantification by S1 analysis

The most important point to keep in mind is that there is no hard and fast regimen for performing the S1 nuclease assay. The procedure given here is but one of several variations and is a modification of the procedure of Quarless and Heinrich (1986). With this procedure it is possible to detect as little as a few hundred femtograms of a specific mRNA, corresponding to 2×10^{-5} % of the poly(A)⁺ component, in 100 μg of total RNA.

1. Wear gloves. Use RNase-free materials, reagents, and technique throughout this procedure. Purified RNA samples are always susceptible to RNase digestion.

Note: Be certain to check the concentration, purity, and integrity of the RNA sample before continuing with this protocol.

2. In a microfuge tube, mix 20–30 μg cellular RNA and sufficient carrier RNA⁶ to give a final mass of 100 μg. Add approximately 1×10^6 cpm of the probe. Adjust the volume of the mixture to 100 μl with sterile H₂O. Be sure to prepare a negative control tube containing carrier RNA and probe only, which should yield no protected fragment.

Note 1: Beginning with 10 μg total cellular or total cytoplasmic RNA is usually sufficient to quantify many mRNAs, though some investigators begin with as much

⁶A stock solution of 1 mg/ml yeast RNA for use as carrier can be prepared in 0.1 M sodium acetate, pH 5.2. Store convenient aliquots at –20°.

as 50 µg of RNA. The quantity of S1 nuclease used in this protocol is more than adequate to completely digest all of the non-hybridized RNA and probe remaining in the reaction.

Note 2: It is critical to ensure that the probe is present in at least a 5- to 10-fold molar excess. In most cases, adding 10^6 cpm of radiolabeled probe will fulfill this requirement.

3. Coprecipitate sample, carrier, and probe with the addition of 11 µl 3M sodium acetate, pH 5.2, and 250 µl ice-cold 95% ethanol. Store at -20° overnight.

Note: The experimental RNA and carrier can be coprecipitated without the probe and stored in hybridization buffer for up to 1 week at -80° .

4. Collect precipitate by centrifugation. Decant supernatant and wash pellet with 70% ethanol (containing 30% sterile H₂O). Discard ethanol (radioactive waste) and invert tube to dry pellet for approximately 10 min. An optional, second wash of the pellet with 95% ethanol will accelerate the drying process. Maintain the RNA pellet slightly damp with ethanol.

Note: Do not Speed Vac or otherwise allow the pellet to dry to completion; doing so will make the RNA virtually impossible to resuspend.

5. Resuspend the pellet in 20 µl S1 hybridization buffer (80% freshly deionized formamide; 40 mM PIPES⁷ pH 6.4; 400 mM NaCl; 1 mM EDTA, pH 8.0). Pipette the sample up and down at least 20 times to ensure complete resuspension.

Note 1: If necessary, the volume of hybridization buffer in which the sample is resuspended may be increased to as much as 30 µl.

Note 2: Hybridization buffer can be prepared in advance and stored in 1 ml aliquots at -80° .

Note 3: If experimental RNA and carrier were coprecipitated ahead of time without the probe (not recommended), add the probe to the pellet first and then add hybridization buffer to a final volume of 30 µl.

Note 4: To detect low abundance mRNAs, it is necessary to keep the RNA concentration near the limit of solubility (5 mg/ml = 100 µg/20 µl) to maximize hybridization kinetics.

Note 5: Hybridization buffer consisting of 80% freshly deionized formamide is a standard formulation. If using a DNA probe, in which case the thermodynamic stability of the hybrid will not be as high as in the case of a cRNA probe, a hybridization buffer consisting of as little as 40–50% formamide may be substituted. Recall that stringency is modulated by temperature (salt and pH, too) and the presence of formamide.

6. Heat the sample in a water bath at 85° for 10 min to fully denature the RNA sample.

Note: It is imperative that the sample be fully denatured, as secondary structures will have an adverse effect on the proper formation of hybrids.

7. Pulse centrifuge the tube(s) for 5 s to collect the hybridization mixture at the bottom of the tube. Immediately place the tube in a prewarmed water bath and incubate the reaction mixture at 50° , or other predetermined temperature, for 3 h.

Note 1: Hybridization temperature is a direct function of the type of probe (RNA or DNA) used, probe length, and G+C content, and must be empirically determined. The typical range for the parameters defined in this protocol is 45 – 60° . In the absence of experience with or information about a particular probe, begin with 50° .

⁷PIPES = Piperazine-1,4-bis(2-ethanesulphonic acid).

Note 2: If carrier RNA was omitted from the reaction tube (step 2), the hybridization interval should be extended to an overnight incubation.

8. Dilute the reaction mixture with the addition of 160 μ l sterile H₂O. Add 20 μ l of 10 \times S1 nuclease buffer that accompanied the purchase of the S1 nuclease. If this pre-made buffer is not available, prepare a generic 10 \times S1 buffer (500 mM sodium acetate, pH 4.5; 2.8 M NaCl; 40 mM ZnSO₄⁸ 50% glycerol).
9. Add sufficient S1 nuclease to produce a final enzyme concentration of 200–400 U/ml. The actual volume required will be variable from one lot to the next, and among vendors.

Note 1: Do not add any more than 10 μ l of S1 nuclease to a 200 μ l digest because the final glycerol concentration will inhibit the reaction. In the case of fewer units of S1 in the reaction it may be necessary to extend the incubation to ensure complete digestion.

Note 2: This part of the protocol, more than any other, is the step that is most likely to require optimization. Recall that AU-rich regions in the hybrid are particularly susceptible. Nicking, if it does occur, will become evident by the presence of one or more unexpected lower molecular weight bands. If this occurs, the remedy is to reduce the amount of enzyme being used to perform the digest.

10. Pulse centrifuge briefly to collect the entire volume in the bottom of the microfuge tube. Incubate at 37° for 1 h.
11. Extract the reaction once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) or phenol:chloroform (1:1) to terminate S1 digestion. Microfuge for 2 min at top speed to separate the phases.
12. Carefully transfer the aqueous (upper) phase to a fresh microfuge tube. Add 1.0 μ l 20 mg/ml glycogen (produces a final glycogen concentration of approximately 100 μ g/ml). Add 2.5 volumes of 95% ethanol to precipitate the hybrids that were protected from nuclease digestion.
13. Incubate tubes on dry ice for 15 min or at –20° for 2 h.
14. Collect the precipitate by centrifugation. Carefully decant the supernatant and allow the pellet to air dry. If desired, it is acceptable to *briefly* dry down the pellet under vacuum, though it is critical not to allow the sample to dry out.
15. Resuspend the pellet in 5 μ l TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).
16. Add 5 μ l of 2 \times formamide loading buffer (50% formamide, 2 \times TBE buffer, 0.2% bromophenol blue). Heat each sample to 95° for 3 min and then cool on ice for 2 min. Electrophorese on a 5% to 6% polyacrylamide/8M urea gel in 1 \times TBE buffer at 250 volts, constant voltage until the dye is close to or at the bottom of the gel. *Alternatively:* Add an equal volume of 2 \times loading buffer (125 mM Tris, pH 6.8; 10% glycerol; 0.2% bromophenol blue) to the sample. Electrophorese on a 5–6% native polyacrylamide gel or use a percentage appropriate for characterizing the size of the expected protected fragment. Use 1 \times TBE (Tris-borate-EDTA) as the electrophoresis buffer. Electrophorese at 250 volts, constant voltage, until the dye front is at or near the bottom of the gel.
17. When electrophoresis is complete, dry down the gel and perform autoradiography at –70° with an intensifying screen (Chapter 14). Alternatively, when high specific activity probes are used, autoradiography can be done on the bench top by plastic wrapping the gel and overlaying a piece of Kodak ReadyPack film.

⁸If the enzyme was supplied in ZnCl₂, then zinc chloride should be substituted in the S1 digestion buffer formulation.

Protocol: transcript quantification by RNase protection

As indicated in the preceding protocol, there are no hard and fast rules for transcript quantification by nuclease protection. The guiding principle is high stringency hybridization, and the RPA offers greater thermodynamic probe: target stability and discriminatory power via the use of RNA probes than is achievable when using DNA probes in conjunction with the S1 nuclease assay. The procedure given here is a modification of the procedures of Zinn *et al.* (1983) and Melton *et al.* (1984) and is but one of several variations, virtually all of which are sold in kit form. This procedure, when optimized, offers sensitivity approaching 0.1 pg of target material, or better. Although the substitution of non-isotopic labeling and detection methods is a possibility, radiolabel continues to offer maximum sensitivity and resolution.

1. Wear gloves. Use RNase-free materials, reagents, and technique throughout this procedure. Purified RNA samples and RNA probes prepared by *in vitro* transcription are always susceptible to RNase digestion.
2. Prepare the antisense probe by *in vitro* transcription and then purify it using any commercially available system or (e.g., Roche Applied Science or Promega) or as described elsewhere (e.g., Gilman, 1987; Sambrook and Russell, 2001).
3. Isolate RNA from cells or tissues under investigation, using any of the protocols described herein.

Note: Be certain to check the concentration, purity, and integrity of the sample before continuing with this protocol.

4. In a microfuge tube, mix 20–30 µg cellular RNA and sufficient carrier RNA to give a final mass of 100 µg. Add approximately 1×10^6 cpm of the probe. Adjust the volume of the mixture to 100 µl with sterile H₂O. Be sure to prepare a negative control tube containing carrier RNA and probe only, which should yield no protected fragment.

Note 1: Beginning with 10 µg total cellular or total cytoplasmic RNA is usually sufficient to quantify many mRNAs, though some investigators begin with as much as 50 µg of RNA. The quantity of RNase used in this protocol is more than adequate to completely digest all of the non-hybridized RNA and probe remaining in the reaction.

Note 2: It is critical to ensure that the probe is present in at least a 5- to 10-fold molar excess. In most cases, adding 10^6 cpm of radiolabeled probe will fulfill this requirement.

5. Coprecipitate sample, carrier, and probe with the addition of 11 µl 3M sodium acetate, pH 5.2 and 250 µl ice-cold 95% ethanol. Store at –20° overnight.

Note: The experimental RNA and carrier can be coprecipitated without the probe and stored in hybridization buffer for up to 1 week at –80°.

6. Collect precipitate by centrifugation. Decant supernatant and wash pellet once with 70% ethanol (containing 30% sterile H₂O). Discard ethanol (radioactive waste) and invert tube to dry pellet for about 10 min. An optional, second wash of the pellet with 95% ethanol will accelerate the drying process. Maintain the RNA pellet slightly damp with ethanol.

Note: Do not Speed Vac or otherwise allow the pellet to dry to completion; doing so will make the RNA virtually impossible to resuspend.

7. Resuspend the pellet in 20 μ l RPA hybridization buffer (80% freshly deionized formamide; 40 mM PIPES, pH 6.4; 400 mM NaCl; 1 mM EDTA, pH 8.0). Pipette the sample up and down at least 20 times to ensure complete resuspension.
Note 1: If necessary, the volume of hybridization buffer in which the sample is resuspended may be increased to as much as 30 μ l.
Note 2: Hybridization buffer can be prepared in advance and stored in 1 ml aliquots at -80° .
Note 3: If experimental RNA and carrier were coprecipitated ahead of time without the probe (not recommended), add the probe to the pellet first and then add hybridization buffer to a final volume of 20 μ l.
Note 4: To detect low abundance mRNAs, it is necessary to keep the RNA concentration near the limit of solubility (5 mg/ml = 100 μ g/20 μ l) to maximize hybridization kinetics.
Note 5: Hybridization buffer consisting of 80% freshly deionized formamide is a standard formulation. Recall that stringency is modulated by temperature (salt and pH, too) and the presence of formamide.
8. Heat the sample in a water bath at 85° for 10 min to fully denature the RNA sample.
Note: It is imperative that the sample be fully denatured, as secondary structures will have an adverse effect on the proper formation of hybrids.
9. Pulse centrifuge the tube(s) for 5 s to collect the hybridization mixture at the bottom of the tube. *Immediately* place the tube in a prewarmed water bath and incubate the reaction mixture at 55° , or other predetermined temperature, for 3 h.
Note 1: Hybridization temperature is a direct function of the probe length and G+C content, and must be empirically determined. The typical range for the parameters defined in this protocol is 45 – 60° . In the absence of experience with or information about a particular RNA probe, begin with 55° .
Note 2: If carrier RNA was omitted from the reaction tube (step 4), extend the hybridization to an overnight incubation.
10. At the conclusion of the hybridization period, dilute the reaction mixture with the addition of 240 μ l sterile H_2O . Add 30 μ l of $10\times$ RNase buffer (100 mM Tris, pH 7.5; 3 M NaOAc, 50 mM EDTA, pH 8.0), followed by the addition of 5 μ l 2 mg/ml RNase A and 5 μ l 10 U/ μ l RNase T1.
Note 1: If more than 20 μ l of RPA hybridization buffer was used in step 7, be sure to increase the volume of H_2O and $10\times$ RNase buffer proportionately.
Note 2: The final concentration of RNase A should be approximately 50 μ g/ml and final concentration of RNase T1 should be approximately 200 U/ml.
11. Pulse centrifuge briefly to collect the entire volume in the bottom of the microfuge tube. Incubate at 37° for 1 h.
12. Extract the reaction once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) or phenol:chloroform (1:1) to terminate RNase digestion. Microfuge samples for 2 min at top speed to separate the phases.
13. Carefully transfer the aqueous (upper) phase to a fresh microfuge tube. Add 1.5 μ l 20 mg/ml glycogen and 2.5 volumes of 95% ethanol to precipitate the hybrids that were protected from nuclease digestion.
14. Incubate tubes on dry ice for 15 min or at -20° for 2 h.
15. Collect the precipitate by centrifugation. Carefully decant the supernatant and allow the pellet to air dry. If desired, it is acceptable to *briefly* dry down the pellet under vacuum, though it is critical not to allow the sample to dry out.
16. Resuspend the pellet in 5 μ l TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

17. Add 5 μ l of 2 \times formamide loading buffer (50% formamide, 2 \times TBE buffer, 0.2% bromophenol blue). Heat each sample to 95° for 3 min and then cool on ice for 2 min. Electrophorese on a 5% to 6% polyacrylamide/8M urea gel in 1 \times TBE buffer at 250 volts, constant voltage until the dye is close to or at the bottom of the gel. *Alternatively:* Add an equal volume of 2 \times loading buffer (125 mM Tris, pH 6.8; 10% glycerol; 0.2% bromophenol blue) to the sample. Electrophorese on a 5% to 6% native polyacrylamide gel or use a percentage appropriate for characterizing the size of the expected protected fragment in 1 \times TBE electrophoresis buffer. Electrophorese at 250 volts, constant voltage, until the dye front is close to or at the bottom of the gel.
18. When electrophoresis is complete, dry down the gel and perform autoradiography at -70° with an intensifying screen (Chapter 14). Alternatively, when high specific activity probes are used, autoradiography can be done on the bench top by plastic wrapping the gel and overlaying a piece of Kodak ReadyPack film.

Troubleshooting

1. Always use freshly prepared probe.
2. Always observe scrupulously the RNase-free strategies described herein and elsewhere.
3. Run a gel to ensure the integrity of the RNA sample just prior to the onset of the assay.
4. Consider lowering the hybridization stringency in the case of no signal, or raising the hybridization stringency if everything “lights up”. The precise temperature calculation will also be dependent on the nature of the probe: recall that double-stranded RNA duplexes are more stable thermodynamically than DNA:RNA duplexes and are therefore able to withstand more stringent hybridization than would the same assay with a DNA probe (discussed in Chapter 13).
5. If no signal is detected and the investigator suspects that the transcript is below the level of detection, perform a quick PCR assay, if primers are available. This will answer the question of whether the transcript is present.
6. Incubation with DNase and RNase, in independent reactions, can be used to demonstrate the specificity of the nuclease protection assay for RNA detection.
7. Reduce the amount sample loaded onto the gel if bands appear to “smile” or if one band obscures visualization of a nearby band because of signal strength.
8. In the case of conflicting data, ensure the probe is present in a substantial (ten-fold) molar excess in each assay. This is performed by molecular titration, in which a constant mass of probe is assayed against various quantities of a particular RNA sample. If the probe is present in excess, changing the starting mass of RNA target should change the signal proportionally.
9. S1 assay is an aggressive enzyme which, in excess concentrations, will attack double-stranded molecules. If difficulties are encountered with this component of the assay, vary the amount of S1 nuclease so that the probe is completely degraded without loss of hybridization signal. If the activity level of a preparation of S1 nuclease is unknown, it may be useful to run an activity titration on a pilot reaction scale before attempting the S1 nuclease assay with valuable experimental material. Lyophilized RNase, as opposed to an RNase cocktail that has been purchased ready for immediate use in the RNase protection assay, generally requires titration and usage optimization.

References

- Ando, T. (1966). A nuclease specific for heat-denatured DNA isolated from a product of *Aspergillus oryzae*. *Biochim. Biophys. Acta* **114**, 158–168.
- Ausserlechner, M.J., R. Sgonc and G. Wick. (1998). Chemiluminescence-based RNase protection assays for simultaneous quantification of procollagen mRNAs containing AU-rich regions. *BioTechniques* **24**(3), 366–368.
- Berk, A.J. and P.A. Sharp. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12**, 721–732.
- Berk, A.J. and P.A. Sharp. (1978). Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci.* **75**, 1274–1278.
- Boorstein, W.R. and E.A. Craig. (1989). Primer extension analysis of RNA. *Methods Enzymol.* **180**, 347–369.
- Chase, J.W. and C.C. Richardson. (1974a). Exonuclease VII of *Escherichia coli*. Purification and properties. *J. Biol. Chem.* **249**, 4545–4552.
- Chase, J.W. and C.C. Richardson. (1974b). Exonuclease VII of *Escherichia coli*. Mechanism of action. *J. Biol. Chem.* **249**, 4553–4561.
- Gilman, M.. (1987). Ribonuclease protection assay. In *Current Protocols in Molecular Biology* (F.M. Ausubel, R. Brent, R.E. Kingston, *et al.*, Eds). John Wiley & Sons, New York. pp. 4.7.1–4.7.8.
- Greenberg, M.E. and E.B. Ziff. (1984). Stimulation of 3T3 cells induces transcription of the *c-fos* proto-oncogene. *Nature* **311**, 433–438.
- Gubler, U. and B.J. Hoffman. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.
- Hofstetter, H., A. Schambock, J. Van Den Berg and C. Weissmann. (1976). Specific excision of the inserted DNA segment from hybrid plasmids constructed by the poly(dA) poly (dT) method. *Biochim. Biophys. Acta* **454**, 587–591.
- Jones, K.A., K.R. Yamamoto and R. Tjian. (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. *Cell* **42**, 559–572.
- Kingston, R.E.. (1988). Primer extension. In *Current Protocols in Molecular Biology* (F.M. Ausubel, R. Brent, R.E. Kingston, *et al.*, Eds). John Wiley & Sons, New York. pp. 4.8.1–4.8.3.
- Kowalski, D., W.D. Kroeker and M. Laskowski. (1976). Mung bean nuclease I. Physical, chemical, and catalytic properties. *Biochemistry* **15**, 4457–4463.
- Kris, R.M., S. Felder, M. Deyholos, G.M. Lambert, J. Hinton, I. Botros, R. Martel, B. Seligmann and D.W. Galbraith. (2007). High-throughput, high-sensitivity analysis of gene expression in Arabidopsis. *Plant Physiol.* **144**, 1256–1266.
- Lee, J.J. and N.A. Costlow. (1987). A molecular titration assay to measure transcript prevalence levels. *Methods Enzymol.* **152**, 633–648.
- Linn, S. and R.J. Roberts. (1982). *Nucleases*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Marzluff, W.F. (1978). Transcription of RNA in isolated nuclei. *Methods Cell Biol.* **19**, 317–332.
- McKnight, S.L. and R. Kingsbury. (1982). Transcriptional control signals of a eukaryotic protein-coding gene. *Science* **217**, 316–324.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn and M.R. Green. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **18**, 7035–7056.

- Quarless, S.A. and G. Heinrich. (1986). The use of complementary RNA and S1 nuclease for the detection of low abundance mRNA transcripts. *BioTechniques* 4(5), 434–438.
- Rosenau, C., B. Kaboord and M.W. Qoronfleh. (2002). Development of a chemiluminescence-based ribonuclease protection assay. *BioTechniques* 33, 1354–1358.
- Sambrook, J. and D.W. Russell (Eds). (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. CSHL Press, Cold Spring Harbor, New York.
- Vogt, V.M. (1973). Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. *Eur. J. Biochem.* 33, 192–200.
- Vogt, V.M. (1980). Purification and properties of S1 nuclease from *Aspergillus*. *Methods Enzymol.* 65, 248–255.
- Weaver, R.F. and C. Weissman. (1979). Mapping of RNA by a modification of the Berk–Sharp procedure: the 5' termini of 15S β -globin mRNA precursor and mature 10S β -globin mRNA have identical map coordinates. *Nucleic Acids Res.* 7, 1175–1193.
- Weigand, R.C., G.N. Godson and C.M. Radding. (1975). Specificity of the S1 nuclease from *Aspergillus oryzae*. *J. Biol. Chem.* 250, 8848–8855.
- Zinn, K., D. DiMaio and T. Maniatis. (1983). Identification of two distinct regulatory regions adjacent to the human β -interferon gene. *Cell* 34, 865–879.

This page intentionally left blank

16 Analysis of Nuclear RNA

Rationale	341	Posthybridization washes and detection	355
Transcription rate assays	343	Protocol: alternative procedure for nuclear run-off assay	356
Relationship to the study of steady-state RNA	346	Protocol: nuclease protection-pulse label transcription assay	357
Nuclear run-off vs. nuclear run-on assay	347	Distinguishing among the activities of RNA polymerases	359
Protocol: nuclear run-off assay	348	Extraction of nuclear RNA for steady-state analysis	360
Harvesting of cells and preparation of nuclei	349	Protocol: direct isolation of nuclear RNA	361
Alternative protocol for preparation of fragile nuclei	349	Protocol: preparation of nuclear RNA from cells enriched in ribonuclease	362
Alternative protocol for preparation of nuclei from whole tissue	350	Troubleshooting nuclear RNA analysis	363
Labeling and recovery of transcripts	351	References	364
Preparation of target DNA	353		
Preparation of RNA for hybridization	354		

Rationale

The modulation of key regulatory molecules is an integral cellular response to both intracellular and extracellular challenge. Among the fundamental goals in the characterization of a biological system are the quantitative assessment of changes in gene expression and the elucidation of the level(s) at which the genes of interest are being controlled. Although potential regulatory points are myriad, they are broadly categorized in eukaryotes as transcriptional or a result of one or more posttranscriptional event(s). Examples of such events include transcriptional (promoter) efficiency, transcription rate, splicing efficiency of precursor hnRNA molecules, nucleocytoplasmic transport, and accessibility of mRNA to the protein translation machinery, to name but a few (Fig. 16.1). The initial characterization of these systems commonly involves isolation, hybridization, and subsequent detection of specific RNA species by Northern analysis, nuclease protection, and PCR-related techniques (Chapters 2, 11, 13, 15, 18, and 19, respectively). While these approaches may furnish excellent qualitative and quantitative

data with respect to steady-state levels of message, i.e., the final accumulation of RNA in the cell, RNA prepared by total cellular lysis provides information neither about the rate of transcription nor about the intracellular location (nuclear or cytoplasmic) of the RNA of interest. Knowledge of this aspect of the cellular biochemistry is important for understanding gene regulation because many mRNAs naturally have much longer half-lives than others and because the half-life of many mRNA species can be modified tremendously in response to a particular xenobiotic regimen or environmental stimulus.

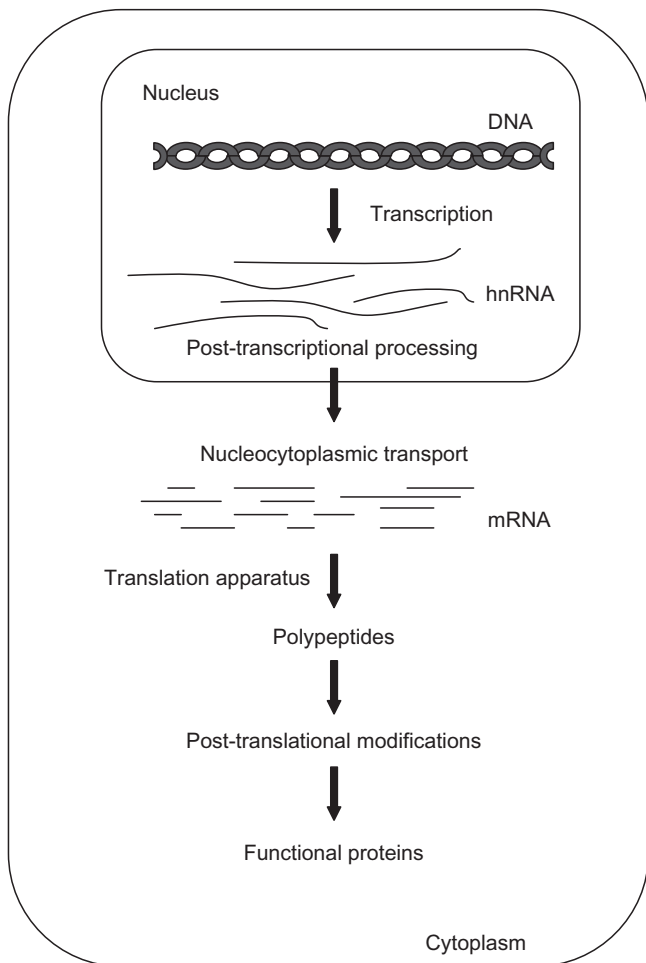


Figure 16.1 mRNA biogenesis. Transcription of genes by RNA polymerase II produces hnRNA, the direct precursor of mRNA. Changes in the cellular biochemistry that affect mRNA maturation are designated as being transcriptional or posttranscriptional in nature, depending on precisely where they exert their influence. Nuclear or cellular lysis yields only those transcripts already accumulated; to assess the relative rates at which transcripts are synthesized, the nuclear run-off assay is utilized.

Transcription rate assays

Two basic approaches have been employed to study the mechanism of transcription of specific genes and the attendant processing of the resulting primary transcripts in eukaryotic cells. In one approach, the rate of transcription is measured in intact nuclei by the incorporation of labeled precursor nucleotides into RNA transcripts which had already been initiated on endogenous chromatin at the time of nuclear isolation. Thus, labeled nuclear RNA is produced by the process of transcriptional elongation. This labeled nuclear RNA is subsequently purified for hybridization to complementary, membrane-bound target DNA sequences. This widely used technique, known as the nuclear run-off assay (Marzluff, 1978; Marzluff and Huang, 1984; Greenberg and Ziff, 1984; Kanazawa *et al.*, 2000) is a very sensitive method for measuring transcription rate as a function of cell state (Fig. 16.2) in both animal and plant cells. The other approach, described below, involves the isolation of steady-state nuclear RNA by Northern analysis, nuclease protection assay, or RT-PCR. A comparison of the traditional Northern analysis, nuclease protection assays, nuclear run-off assay, and the polymerase chain reaction (RT-PCR) is presented in Table 16.1.

The principal advantage of the nuclear run-off assay is that labeling occurs while maintaining the natural geometry of the transcription apparatus. Because the nuclear run-off assay yields unspliced precursor RNA, however, the investigator is uncertain as to the natural posttranscriptional fate of these RNA

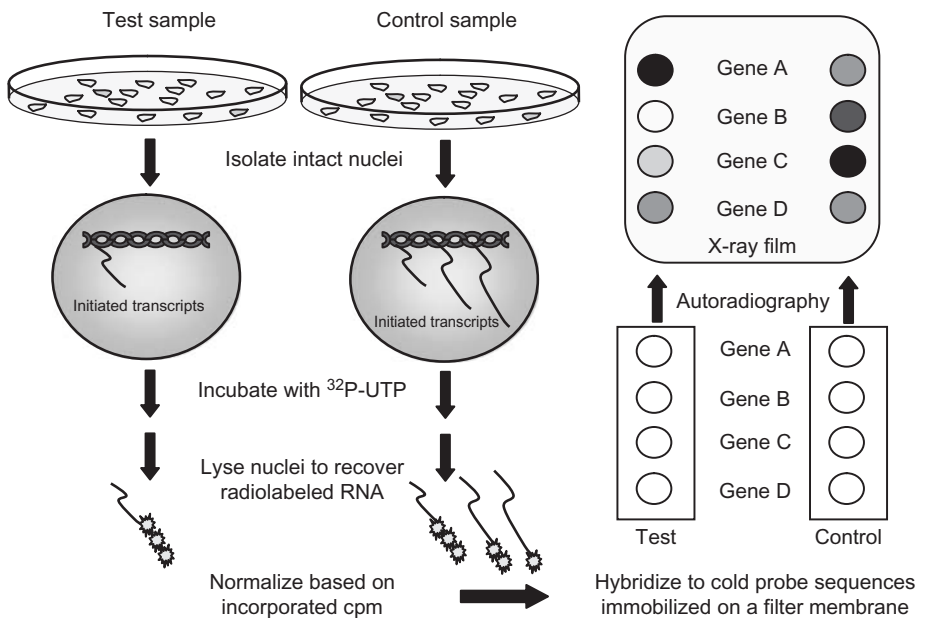


Figure 16.2 Nuclear run-off assay. Relative rate of transcription of all genes can be assessed by incubation of intact nuclei with an NTP cocktail containing labeled UTP.

Table 16.1 Comparison of Northern Analysis, Nuclease Protection Assay, Nuclear Run-off Assay, and RT-PCR

	Northern Analysis	Nuclease Protection Assay	Nuclear Run-off Assay	RT-PCR
Advantages	<p>Provides a qualitative component to RNA analysis.</p> <p>Supports several rounds of hybridization with different probes.</p> <p>Is compatible with total, cytoplasmic, nuclear or poly(A)⁺ RNA.</p> <p>RNA is relatively stable on filter.</p> <p>Is able to assess integrity of the sample.</p>	<p>Much higher sensitivity than Northern analysis.</p> <p>Requires less handling of RNA than other types of analysis.</p> <p>Is tolerant of partially degraded RNA.</p> <p>Solution hybridization is more quantitative than filter hybridization.</p> <p>Can be used for steady-state or transcription rate assays.</p>	<p>Characterizes relative rate of transcription.</p> <p>Natural geometry of the chromatin is maintained.</p> <p>Permits study of several genes simultaneously.</p> <p>Can be used to discern transcriptional vs. posttranscriptional gene regulation when used in conjunction with data from Northern analysis.</p>	<p>Provides unparalleled sensitivity when properly designed.</p> <p>Requires very small amounts of starting material.</p> <p>Unparalleled resolution.</p> <p>Supersedes many of the classical techniques.</p> <p>Minimizes the amount of handling of the RNA.</p> <p>Very rapid technique, favoring productivity.</p>
Disadvantages	<p>Is the least sensitive assay.</p> <p>Denaturants can be toxic.</p> <p>Requires extensive handling of RNA.</p> <p>Is a time-consuming process.</p> <p>Provides ample opportunity for RNase degradation.</p> <p>Characterizes only steady-state RNA.</p>	<p>Protected fragment is smaller than native RNA.</p> <p>Nucleases, especially S1, can be difficult to control.</p> <p>Assay is more sensitive to exact hybridization parameters than other assays.</p> <p>Double-stranded probes can compromise sensitivity of the assay if reannealing occurs.</p>	<p>Nuclear isolation requires a fair amount of skill.</p> <p>Probe complexity is very large.</p> <p>Unlabeled endogenous RNA can compete with labeled RNA during hybridization.</p> <p>Mechanics of the assay support transcript elongation, and not initiation, during labeling.</p>	<p>Much more sensitive to the precise reaction components and conditions than the other assays.</p> <p>Exquisitely sensitive to contaminants, especially genomic DNA.</p> <p>Carryover contamination must be addressed.</p> <p>Optimization can be time-consuming and costly.</p>

molecules. The mechanics and reaction conditions of the nuclear run-off assay promote the elongation of initiated transcripts but are not believed to support new initiation events¹. The degree of labeling of any particular RNA species (and hence, the relative transcription rate of specific genomic sequences) may then be assessed by liquid scintillation counting coupled with autoradiography (Fig. 16.3). Alternatively, this procedure can be carried out using non-radioactive variants of the some of the established procedures (Patrone *et al.*, 2000; Kwei *et al.*, 2004). Under a defined set of experimental conditions, these data correlate directly with the number of RNA polymerase molecules engaged in transcribing a specific gene and indirectly with the transcriptional efficiency of regulatory sequences associated with the gene(s) of interest.

A key advantage of the nuclear run-off assay is the ability to simultaneous assay transcriptional activity at multiple genomic loci, all of which are presumably transcribed in the same relative amounts in isolated nuclei as in intact cells. Because a “typical” eukaryotic cell is actively transcribing thousands of different genes at any given point in the cell cycle, the complexity² of the labeled RNA (which acts as a heterogeneous probe in this assay) is enormous. Consequently, the concentration of the transcribed RNA sequences of immediate interest in the experiment may be only a very small percentage of the mass of the probe. Despite this, the sensitivity of the assay is excellent because (1) the labeled RNA is single stranded, so a reduction of effective probe concentration due to strand renaturation is a non-issue; (2) the RNA is labeled continuously along its length to very high specific activity (Chapter 12); and (3) the thermodynamic stability of RNA:RNA hybrids is greater than RNA:DNA

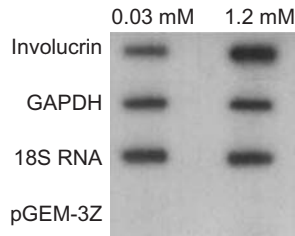


Figure 16.3 Nuclear run-off assay using normal human keratinocytes (NHK). Cells were cultured in medium supplemented with 0.03 or 1.2 mM calcium in the presence of ³²P-UTP. The newly synthesized RNA was then hybridized to pGEM-3Z containing involucrin cDNA, GAPDH cDNA, and cDNA of 18S RNA that was blotted onto a filter. pGEM-3Z was used as the negative control. Ng, D.C. *et al.* 1996. *Frontiers in Bioscience* 1, 16–24 (with permission).

¹Large amounts of unlabeled, endogenous nuclear RNA can competitively interfere with an assay designed to measure newly synthesized RNA, one aspect of nuclear biochemistry that this assay does not support. Details for quantifying transcript initiation have been described elsewhere (Marzluff and Huang, 1984; pp. 103–119).

²Complexity in this context refers to the total length of *different* RNA sequences acting as probe.

and DNA:DNA duplexes. By tolerating very stringent posthybridization washes, the signal-to-noise ratio is greatly improved.

Relationship to the study of steady-state RNA

Cellular lysis and subsequent RNA isolation performed by any of the methods described in Chapters 2–5 renders what is known as steady-state RNA. One might consider steady-state RNA to be the final accumulation of RNA in the cell at the moment of cell lysis. The precise composition of steady-state RNA is function of the rate of synthesis as well as the rate of turnover (degradation) in the cell under a defined set of conditions. One might consider this mnemonic:

$$\text{relative synthesis} + \text{relative degradation} = \text{relative abundance}$$

Consequently, a biochemical snap-shot of the abundance of a particular RNA transcript is not indicative of the rate of transcription at the associated gene locus, the age of the transcript in the cell, or the rate at which identical transcripts are being degraded. By measuring the rate of gene transcription and in comparison with steady-state data, for example by Northern analysis with cytoplasmic RNA, it is possible to assign observed modulation of gene expression to the transcription level (up- or downregulation) or to a posttranscriptional event. See Table 16.2 for an example of how such data might be presented. Best of all, the mechanics of transcription rate assays maintain the natural geometry of the chromatin so that the rate data mirrors the *in vivo* conditions as closely as possible.

The nuclear run-off assay consists of several distinct phases, each of which has a distinct goal (Table 16.3). The degree of success of the nuclear run-off assay is almost entirely dependent on the speed with which nuclei can be isolated from intact cells and radiolabeled with precursor UTP. The most critical parameter by far is the preparation of nuclei before the labeling step. Failure to generate high specific activity RNA probe is usually a direct result of inexperienced handling of the nuclei prior to the labeling step. Nuclear isolation must also be carried out in such a way as to preserve RNA polymerase activity and

Table 16.2 Dependence of Cellular Oncogenes on Cell Shape

Cell line	Proliferation phenotype	Contact inhibited	Relative transcription rate			
			<i>c-ras</i>	<i>c-myc</i>	<i>c-fos</i>	<i>c-fes</i>
CUA-1	Shape dependent	+	30.0	16.0	0.92	0.93
HT-1080	Shape independent	–	0.86	0.67	0.93	0.92
HT-IFN ^R	Shape dependent	+	0.38	0.51	0.64	0.66

Values reflect the abundance of gene-specific nuclear transcripts in flat cells divided by the transcript abundance in rounded cells plated onto the ethanol-soluble polymer poly[HEMA]. Abundance levels are determined from image analysis performed on nuclear run-off autoradiograms (Farrell and Greene, 1992).

nuclear structure during the isolation from cells cultured *in vitro* or from tissue (Ernest *et al.*, 1976). Most protocols appearing in the literature vary with respect to the postlabeling RNA purification steps only. Methods for the isolation of chromatin which retain endogenous RNA polymerase activity have also been described (Marzluff and Huang, 1975).

In general, nuclei are isolated from whole cells by incubation in hypotonic lysis buffer featuring the inclusion of the gentle, non-ionic detergent NP-40 (Price and Penman, 1972; Greenberg, 1988); in isosmotic sucrose buffer containing Triton X-100 (Marzluff *et al.*, 1973; 1974) or by non-aqueous methods³ (Gurney and Foster, 1977; Lund and Paine, 1990). Nuclei may then be labeled immediately or stored frozen in cell freezing ampoules (*never* microfuge tubes) in liquid nitrogen⁴ or at -80° for several months without significant loss of labeling potential.

Nuclear run-off vs. nuclear run-on assay

A fair amount of confusion exists over the nomenclature associated with the study of *in situ* transcription. The terms nuclear run-on assay and the nuclear run-off assay technically refer to two different approaches, though these two terms have long been used interchangeably. The original intent of the term “nuclear run-off” was to describe those *in vitro* methods by which the individual

Table 16.3 Major Steps and Goals Associated with Nuclear Transcription Assays

Procedural Phase	Goal
Nuclear isolation	Rapid isolation of intact nuclei capable of supporting transcript elongation.
Transcript elongation	Incorporate labeled NTP precursors into nascent transcripts while maintaining the natural geometry of the transcriptional apparatus.
Recovery of labeled transcripts	Removal of protein and DNA from the newly labeled RNA.
Hybridization of labeled transcripts to specific membrane-bound sequences	Hybridize labeled transcripts to their cognate cDNA. There is no limit to the number of sequences that can be assayed in this manner.
Quantification of hybridization events	Determine the relative transcription rates by autoradiography and/or by scintillation counting. The hybridization signal correlates with the degree to which RNAs of interest are present in the pool of all transcribed, labeled RNAs.

³Non-aqueous methods of nuclear isolation consist of tissue/cell lyophilization, homogenization of the dry powder in glycerol (for example), and then nuclear sedimentation in non-aqueous buffer.

⁴If unfamiliar with methods for the safe handling of liquid nitrogen, liquid nitrogen Dewars, and the freezing and thawing of tubes, be sure to seek the guidance of someone who is experienced in this area. This is necessary for the safety of everyone in the lab.

factors required to support transcription where elucidated, one by one, by adding various components to a DNA template, often a restriction fragment, much as in the manner of working with nuclear- or whole-cell extracts. When conditions supporting transcriptional activation are met, the RNA polymerase will “run-off” the end of the template, producing transcripts of uniform length. The quantity of RNA produced is therefore an indicator of the efficiency of transcription.

The term “nuclear run-on”, in the earliest days of development of transcription assays, was intended to refer to methods for the identification of genes being expressed at a particular moment under a defined set of conditions. One would rapidly isolate nuclei from cells of interest and, in the presence of labeled nucleotides, promote short-term elongation and concomitant radiolabel incorporation of initiated transcripts.

Sometime in the 1980s the terms “nuclear run-off assay” and “nuclear run-on assay” started to be used interchangeably. Many peer-reviewed papers were published describing the labeling of RNA in isolated nuclei; some authors called the technology a run-off assay while others referred to the same strategy as a run-on assay. It now seems that these two terms are taken to mean the same thing, namely the elongation and labeling of initiated transcripts in nuclei isolated from cells and tissues and, to an extent, the usage of one term over another seems to be regional in nature. Inasmuch as “nuclear run-off assay” is far more commonplace, this nomenclature will be used throughout this laboratory guide to refer to the transcription rate assays which follow.

Protocol: nuclear run-off assay⁵

In this laboratory, the best labeling has occurred when cells are maintained as cold as possible until the initiation of the labeling reaction. It is very helpful to fill two or three large trays (such as those often used for autoclaving bottles) with ice so that tissue culture vessels can be set directly on ice for the first three steps in this protocol. This precludes the unpleasantness of having to work in a cold room. Further, harvesting cells from tissue culture is greatly accelerated when cells are grown in 100- or 150mm tissue culture dishes rather than in tissue culture flasks.

Cells may be harvested by trypsinization (Appendix J). We have learned the hard way that overtrypsinization greatly reduces the labeling efficiency in this assay. While the synthesis of hnRNA in “quasi-transformed” cell populations appears to be unaffected by the cell shape distortions that undoubtedly accompany trypsinization (Ben-Ze’ev *et al.*, 1980), RNA synthesis in diploid cells shows a much greater shape-responsiveness to even subtle changes in morphology (Benecke *et al.*, 1978; Wittelsberger *et al.*, 1981). For this reason, scraping cells from dishes resting on ice is the method least likely to cause distortion of the cellular morphology and is therefore strongly recommended.

⁵The protocols presented here are improvements on the procedures of Groudine *et al.* (1981), Nevins (1987), and Greenberg (1988).

Harvesting of cells and preparation of nuclei

1. Decant and discard cell growth medium from adherent cells growing in tissue culture.

Note: While as few as 5×10^6 nuclei may be utilized in a labeling reaction, it is wise to begin with 3 to 5×10^7 nuclei to achieve efficient labeling. Sufficient counts per minute (cpm) must be generated to titrate the RNA sample at the hybridization step. Cell number can be determined quickly by hemocytometer counting just prior to lysis.

2. Wash the monolayer once with 10 to 15 ml ice cold phosphate-buffered saline (PBS). Decant most of the PBS, leaving approximately 1 ml in the tissue culture vessel.

3. Using a sterile cell scraper, harvest the cells by gentle but rapid scraping.

Note: Collect cells grown in suspension by centrifugation at $500 \times g$ for 5 min at 4° in a precooled rotor. Decant and discard the supernatant. Then resuspend the cell pellet in 10 ml of ice-cold $1 \times$ PBS and pellet as described above. Remove the supernatant as completely as possible and then proceed to step 5.

4. Transfer the cells to a pre-chilled centrifuge tube (e.g., a 15- or 30 ml nuclease-free Corex glass tube) and collect cells by centrifugation at $500 \times g$ for 5 min at 4° . Remove the supernatant as completely as possible.

Note 1: Transport tubes containing cells and/or nuclei to and from the centrifuge in an ice bucket.

Note 2: Make sure that glass centrifuge tubes are supported by the proper rubber adapter sleeves when placed into the centrifuge rotor.

5. Prior to resuspending the cells, loosen the pellet by very gentle vortexing for 5 s, which will prevent clumping of cells. Slowly add 4 ml of hypotonic lysis buffer (10 mM Tris, pH 7.4; 10 mM NaCl; 5 mM $MgCl_2$; 0.5% NP-40) while continuing to vortex gently. Incubate on ice for 5 min.

6. Collect the nuclei by centrifugation at $500 \times g$ for 5 min at 4° .

Note: If desired, cytoplasmic RNA can be purified from the resulting cytoplasmic supernatant for Northern analysis. Store supernatant on ice and continue working up nuclei first.

7. Resuspend the nuclear pellet in 4 ml of the lysis buffer described in step 5. Collect the nuclei by centrifugation at $500 \times g$ at 4° .

8. Carefully decant and discard the supernatant. Resuspend the nuclei in 200 to 300 μ l of prechilled glycerol storage buffer (50 mM Tris, pH 8.0; 30% glycerol; 2 mM $MgCl_2$; 0.1 mM EDTA) by gentle vortexing as described in step 5. The nuclei are now ready for immediate labeling or may be stored in sealed cell freezing ampoules in liquid nitrogen or at -80° for several months.

Alternative protocol for preparation of fragile nuclei⁶

In this procedure, fragile nuclei are prepared by cell lysis in isosmotic sucrose buffer, and pelleted into a dense cushion of sucrose. This approach helps maintain nuclear integrity by preventing nuclei from slamming into the bottom of the centrifuge tube. Inasmuch as the density of nuclei varies from one cell type to the next, tweaking the buffer formulations to accommodate nuclear purification from a specific cell type may be necessary. As in the previous protocol, the recovery of cells from tissue culture is greatly accelerated when cells are grown in 100- or 150 mm tissue culture dishes rather than in tissue culture flasks.

⁶This is a modification of the procedure of Marzluff and Huang (1984).

1. Decant and discard cell growth medium from adherent cells growing in tissue culture.
Note: Begin with 3 to 5×10^7 nuclei to generate sufficient cpm to titrate the RNA sample at the hybridization step. Cell number can be determined quickly by hemocytometer counting just prior to lysis.
2. Wash the monolayer once with 10 to 15 ml cold PBS. Decant most of the PBS, leaving approximately 1 ml in the tissue culture vessel.
3. Using a sterile cell scraper, harvest the cells by gentle but rapid scraping.
Note 1: Cells may be harvested by trypsinization (Appendix J).
Note 2: Collect cells grown in suspension by centrifugation at $500 \times g$ for 5 min at ambient temperature. Decant and discard the supernatant. Optional: wash the cell pellet one time with an aliquot of cold PBS and pellet as described earlier. Remove the supernatant as completely as possible and then proceed to step 5.
4. Transfer the cells to a suitable centrifuge tube (e.g., 15- or 30 ml nuclease-free Corex glass tube) and collect cells by centrifugation at $500 \times g$ for 5 min at room temperature. Remove the supernatant as completely as possible.
Note 1: Transport tubes containing cells and/or nuclei to and from the centrifuge in an ice bucket.
Note 2: Make sure that glass centrifuge tubes are supported by the proper rubber adapter sleeves when placed into the centrifuge rotor.
5. Resuspend the cells in prechilled sucrose buffer (0.32M sucrose; 5 mM CaCl_2 ; 3 mM magnesium acetate; 0.1 mM EDTA; 0.1% Triton X-100; 1 mM DTT; 10 mM Tris, pH 8.0) at a density of 1×10^7 cells/ml.
Note: This buffer can be prepared in advance; however, do not add DTT or Triton X-100 until just prior to use.
6. Homogenize the cells in a Dounce homogenizer with a maximum of 10 strokes of a tightly-fitting pestle.
Note: One common cause for poor label incorporation is the use of a loosely fitting pestle and the concomitant failure to liberate a large number of nuclei.
7. Fill the bottom one-third of an RNase-free ultracentrifuge tube with an aliquot of cushion buffer (2M sucrose; 3 mM magnesium acetate; 1 mM DTT; 10 mM Tris, pH 8.0). Dilute the homogenate with an appropriate aliquot of cushion buffer to produce a sufficient volume to occupy the remaining two-thirds of the ultracentrifuge tubes. Carefully layer the resulting mixture over the sucrose cushion already in place.
8. Collect the nuclei by centrifugation at $30,000 \times g$ for 45 min at 4° .
9. Resuspend the nuclear pellet in storage buffer (40% glycerol; 5 mM magnesium acetate; 0.1 mM EDTA; 5 mM DTT; 50 mM Tris, pH 8.0) at a density of 10^8 nuclei/ml.
10. Immediately seal nuclei in cell freezing ampoules and store in liquid nitrogen or at -80° for several months. Otherwise, initiate labeling protocol without delay.

Alternative protocol for preparation of nuclei from whole tissue⁷

1. Harvest tissue according to standard procedures. Rapidly mince tissue and rinse with an ice cold solution of 0.14 mM NaCl; 10 mM Tris, pH 8.0.
2. Homogenize minced tissue in cold lysis buffer (0.32M sucrose; 5 mM CaCl_2 ; 3 mM magnesium acetate; 0.1 mM EDTA; 0.1% Triton X-100; 1 mM DTT; 10 mM Tris, pH 8.0). Use 5 ml lysis buffer per gram of tissue.

⁷This is a modification of the procedures of Derman *et al.* (1981) and Marzluff and Huang (1984).

3. *Rapidly* filter homogenate through several layers of sterile cheesecloth. This will facilitate removal of connective tissue and other debris.
4. Dounce homogenize filtrate with 10–20 strokes of a tightly fitting pestle. Remove a 5 μ l aliquot of the homogenate and place on a microscope slide, overlay with a coverslip and assess the extent of cellular lysis by light microscopy.
5. Add 1 to 2 volumes of prechilled solution of 2.0M sucrose⁸; 3 mM magnesium acetate; 0.1 mM EDTA; 1 mM DTT; 10 mM Tris, pH 8.0 to the homogenate. Layer the resulting mixture over a cushion of this same buffer that occupies one-third of the volume of an RNase-free ultracentrifuge tube.
6. Collect nuclei by centrifugation at 30,000 \times g for 45 min at 4°.
7. Resuspend the nuclear pellet in glycerol storage buffer (25% glycerol; 5 mM magnesium acetate; 0.1 mM EDTA; 5 mM DTT; 50 mM Tris, pH 8.0) at a density of 1×10^8 nuclei/ml.
8. Immediately seal aliquots of nuclei in cell freezing ampoules and freeze in liquid nitrogen or at -80° for several months. Otherwise, initiate labeling protocol without delay.

Labeling and recovery of transcripts

Note: It is very important to have the 2 \times reaction buffer prepared before thawing nuclei. Permitting thawed nuclei to sit for just a few minutes on ice will greatly diminish labeling efficiency.

1. Prepare a 2 \times reaction buffer (10 mM Tris, pH 8.0; 5 mM MgCl₂; 300 mM KCl; 1 mM ATP, 1 mM CTP, 1 mM GTP; 5 mM DTT).

Note: Using an NTP stock solution from any of several vendors will give the greatest reproducibility from one experiment to the next. Because NTP stock solutions prepared in the lab are lacking in quality control, so to speak, the labeling efficiency is likely to be impacted.

2. Add 200 μ l of the 2 \times reaction buffer to each 200 μ l aliquot of nuclei. If nuclei were previously stored frozen, add 200 μ l of the 2 \times reaction buffer to 200 μ l of thawed nuclei.

Note 1: In this laboratory, optimum labeling efficiency is observed when 200 μ l of 2 \times reaction buffer is added to the ampoule containing the frozen nuclei immediately after placing it on ice to thaw. Do not wash nuclei after thawing, and thaw an aliquot only once.

Note 2: Do not be alarmed if the mixture of thawed nuclei appears viscous. Although frozen and thawed nuclei may lyse to some extent, transcriptional complexes remain intact and are competent to support the elongation of initiated transcripts (Marzluff and Huang, 1975; Nevins, 1987). Of course, care should be taken in each step prior to the labeling reaction to minimize the amount of nuclear damage.

3. Immediately add 10 μ l [α -³²P] UTP 800 Ci/mmol, 10 mCi/ml, and incubate for 30 min at 26° with gentle shaking.

Note 1: Labeling can be carried out directly in the freezing ampoule. Alternatively, freshly prepared nuclei are best labeled in a larger polypropylene tube to better manage the spread of radioactivity that will invariably result from the mechanics of this assay.

⁸The exact sucrose concentration is a function of the cell type under investigation and should be empirically determined before using valuable starting material. The usual range is 1.8–2.2M.

Note 2: In this laboratory, the most reproducible degree of label incorporation was observed when the labeling reaction is shaken at 30–50rpm on a temperature-controlled orbital platform (Lab-Line Instruments).

4. In a separate, RNase-free tube, mix 1 ml of high salt buffer (500 mM NaCl; 50 mM MgCl₂; 2.5 mM CaCl₂; 10 mM Tris, pH 7.4) with 50 μl of 1 mg/ml RNase-free DNase I, just prior to use. Add 750 μl of the resulting mixture to each tube containing labeled nuclei.

Note: The high salt buffer will terminate the labeling reaction and disrupt all of the remaining intact nuclei. The nuclear disruption greatly increases the viscosity of the solution, because of the liberation of chromatin. Use a silanized⁹, RNase-free Pasteur pipette, a micropipettor, or a 25-gauge needle attached to a 1 ml tuberculin syringe to shear the DNA several times.

5. Incubate this mixture for 5 min at 37°.

Note: The purpose of this brief treatment with DNase I is to partially fragment, rather than completely degrade, genomic DNA so that the solution is physically more manageable and less likely to trap labeled RNA molecules in the subsequent purification steps. Partially degraded DNA is removed in subsequent steps.

6. Add 200 μl SDS/Tris buffer (5% SDS; 500 mM Tris, pH 7.4; 125 mM EDTA) and 15 μl of 20 mg/ml proteinase K. Incubate for 30 min at 42° to ensure extensive protein hydrolysis.
7. Extract each sample with 1.25 ml of phenol:chloroform:isoamyl alcohol (25:24:1). Centrifuge for 8–10 min at 800 × g to separate the phases.
8. Transfer the upper aqueous phase to a fresh tube and add, in order:
2 ml sterile (nuclease-free) H₂O
3 ml 10% TCA¹⁰ in 60 mM sodium pyrophosphate
15 μl 10 mg/ml carrier RNA (e.g., junk yeast RNA).

Incubate on ice for 45 min.

Note: These steps will result in precipitation of the labeled RNA.

9. Filter the precipitate onto Whatman GF/A glass fiber filters using a Millipore Sampling Manifold (Millipore, Billerica, MA), occasionally referred to as a vacuum filtration bank. Wash each filter three times with 10 ml aliquots of 5% TCA, 30 mM sodium pyrophosphate.

Note: Omission of the TCA precipitation step and filtration will usually result in unacceptably high levels of background even after the most stringent posthybridization washes. This is a key step contributing to the quantitiveness of this assay.

10. Transfer each filter to a glass scintillation vial and add 1.5 ml DNase I buffer (20 mM HEPES (free acid), pH 7.5; 5 mM MgCl₂; 2.5 mM CaCl₂). Add 50 μl 1 mg/ml RNase-free DNase I and incubate for 30 min at 37°.
11. Terminate the digestion with the addition of 40 μl of 500 mM EDTA, pH 8.0, and 70 μl of 20% SDS.
12. Elute the purified RNA from each filter paper by heating the samples to 65° for 10 min. *Carefully* transfer all of the supernatant to a fresh tube.
13. Repeat the elution step by adding 1.5 ml of elution buffer (1% SDS; 10 mM Tris, pH 7.5; 5 mM EDTA) to the scintillation vials containing the filters. Heat to 65° for 10 min and then *carefully* recover the supernatant and combine it with the original supernatant from step 9.

⁹See Appendix I for hints on silanizing glassware.

¹⁰TCA = trichloroacetic acid.

14. Extract the resulting RNA solution one time with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), followed by a single extraction with an equal volume of chloroform.

Note: In both instances, the RNA will partition to the aqueous (upper) phase.

15. Transfer the aqueous material (approximately 3 ml) to a silanized, RNase-free Corex glass tube (or the equivalent).

16. Optional step: Add 650 μ l of 1 N NaOH and incubate on ice for no more than 3 min and then neutralize the reaction with the addition of 1.5 ml 1 M HEPES (free acid), pH 7.4.

Note: The limited hydrolysis of RNA tends to improve hybridization efficiency (Jelinek et al., 1974). Incubation for more than 3 min runs the great risk of RNA degradation by alkaline hydrolysis. Further, limited hydrolysis reduces the likelihood of high-molecular-weight RNA self-hybridization, in favor of forward, intermolecular hybridization kinetics. If desired, this step and the neutralization step that follows can be eliminated altogether.

17. Precipitate the RNA with the addition of 0.1 volume of 3 M sodium acetate, pH 5.2 (approximately 500 μ l) and 2.5 volumes of ice cold 95% ethanol (approximately 15 ml). Store overnight at -20° .

Preparation of target DNA

DNA sequences corresponding to the genes of interest are best prepared by PCR. In the event that the genes of interest have been cloned into plasmids, the best strategy is to use PCR primers for the immediately-flanking vector sequences. This precludes the need to excise the cDNA by restriction endonuclease digestion and concomitant gel purification of the liberated insert.

The nuclear run-off assay can be used to simultaneously assess the extent of transcription of a variety of genes, and it is important to include in this assay at least one gene with a purported "housekeeping" function, that is, a gene whose level of expression is not expected to vary under the defined experimental conditions. One worthwhile blotting strategy is to apply the target DNAs corresponding to the different genes of interest in the same vertical column of the blotting manifold so that the filter can be cut into narrow strips and hybridized to the extremely heterogeneous RNA probe in a minimum volume of hybridization buffer.

If using a nylon membrane, retention of these target DNA sequences is a non-issue; however, the target DNA should be at least 1.0 kb to ensure quantitative retention if using nitrocellulose. In this regard, and in consideration of its poor binding capacity, it is wise to avoid nitrocellulose altogether. Plan to load 2 to 5 μ g of denatured DNA to each well in a volume of 100 to 200 μ l.

1. Denature linear target DNA with the addition of 0.1 volume of 1 N NaOH. Incubate at room temperature for 30 min or at 37° for 10 min.
2. Neutralize the alkaline pH of the mixture by the addition of 10 volumes of ice cold $6\times$ SSC.
3. Store denatured DNA on ice until ready to blot, but for no longer than 15–20 min. It is best to complete blotting within 15 min after the addition of $6\times$ SSC to preclude the reannealing of the denatured sample.

4. Assemble a Minifold I (dot blot) or Minifold II (slot blot) apparatus (Whatman, Ltd.) and begin to apply a low vacuum. Pipette 2 to 5 μg (or other predetermined mass) of denatured DNA to each well in a volume of 100 to 200 μl . After the buffer has been pulled through the filter, wash each well with 300 μl 6 \times SSC.

Note 1: Be sure to load control wells that contain (1) a corresponding mass of linear, non-recombinant vector DNA, or other carrier DNA; and (2) buffer only, in order to assess non-specific binding to non-homologous DNA and to the filter itself, respectively.

Note 2: The Slot Blot™ apparatus will concentrate the sample into a smaller area (6 mm²) than the Dot Blot™ apparatus (12.5 mm²); thus slot-blotting is expected to yield data that are easier to quantify than dot-blotting because the slot configuration permits easier quantitation with image analysis software.

5. Immobilize the target DNA onto the filter paper for subsequent hybridization according to the instructions of the manufacturer. DNA may be cross-linked to nylon membranes using a calibrated UV light source while nitrocellulose is baked for 2 h at 80° *in vacuo*. Alternatively, nylon filters may be baked instead, for as little as 1 h at temperatures as low as 65° (vacuum not required). The filter may then be used immediately or stored in a cool, dry location for later use. Details pertaining to nucleic acid immobilization on solid supports can be found in Chapter 11.

Preparation of RNA for hybridization

1. Recover the precipitated RNA by centrifugation at 10,000 $\times g$ for 30 min at 4°. Carefully decant as much of the ethanol as possible into radioactive waste. Allow the tubes to drain briefly, and then wash once with 1 ml 70% ethanol (dilute 95% ethanol with sterile H₂O). Centrifuge, if necessary, and then decant as much of this ethanol wash as possible. Allow the tubes to drain briefly. If excess ethanol remains, carefully wash the RNA once with 1 ml 95% ethanol, centrifuge if necessary, and allow the tube(s) to drain briefly.
2. Resuspend the RNA pellet in 1 ml of TES¹¹ solution (10 mM TES, pH 7.4; 10 mM EDTA; 0.2% SDS). Shake the sample at 100 rpm on an orbital platform for 30 min at room temperature to redissolve the RNA.
3. Remove a 5 μl aliquot from each sample and count for 1 min to determine cpm/ml. Normalize the samples with respect to cpm/ml by the addition of TES solution, to provide nearly equal counts per hybridization. In general, the cpm/ml averages about 10⁷ cpm/ml per 5 \times 10⁷ nuclei when the labeling reaction is efficient.

Note: The normalization of samples by cpm is based on the assumption that the overall level of RNA synthesis is not changing as a function of cell state, but rather that specific genes are being up- or downregulated. The investigator must empirically determine the reasonableness of this assumption for each new set of cell state conditions and also ensure equal numbers of nuclei at the onset of the experiment.

4. Cut the filter containing the immobilized target DNA sequences into strips such that each strip contains one complete set of target DNAs to be hybridized to the labeled RNA probes.
5. Hybridization is most efficiently and reproducibly executed directly in a clean 5 ml scintillation vial. This may be accomplished by coiling the filters strips containing the DNA and carefully pushing each strip to the bottom of individual scintillation

¹¹TES = [N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid].

vials such that the side of the filter onto which the DNA is fixed is facing the center of the tube.

Note 1: Hybridization in this fashion will minimize the amount of probe required and will maintain the maximum probe concentration. Moreover, this approach will minimize the radioactive mess that is invariably associated with hybridization in plastic bags.

Note 2: To ensure that the scintillation vial is RNase-free it is strongly suggested that the vial be soaked in H_2O_2 for a minimum of 20 min, and then rinsed thoroughly with sterile H_2O .

Note 3: Pre-wet the filter strips in sterile H_2O and briefly blot away excess water (Whatman 3MM paper) before placing it into the scintillation vial for hybridization.

6. Mix 1 ml of RNA solution with 1 ml of TES/NaCl (10 mM TES, pH 7.4; 10 mM EDTA; 0.2% SDS; 600 mM NaCl). Add the resulting 2 ml of RNA probe solution to the scintillation vials containing the target DNA filter strips, close the tubes tightly, and hybridize for 24–36 h at 65°. Be certain that filters in the vials are completely immersed in hybridization solution.

Note: In this lab, the most efficient method for hybridization is to insert individual scintillation vials into the disposable polystyrene racks in which conical 15 ml centrifuge tubes are often packaged. The rack is then placed on the platform of an orbital shaker incubator preheated to 65° and agitated slowly (50–75 rpm) for the duration of the hybridization.

Posthybridization washes and detection

1. At the conclusion of the hybridization period, wash the filters in batch for 1 h in an excess volume of 2× SSC, preheated to 65°.
2. Place filters in individual glass scintillation vials containing 8 ml of 2× SSC and 12 μl of 10 mg/ml RNase A. Alternatively, non-specific RNA molecules can be digested with a combination of RNase A (5.0 μg/ml) and RNase T1 (5 U/ml) in 2× SSC. Incubate at 37° for 30 min without shaking.
3. Wash the filters in batch for 1 h in an excess volume of 2× SSC at 37°. Briefly blot the filters, ensuring that they do not dry out completely. Plastic-wrap the damp filters and then tape them to a solid support such as Whatman 3MM paper or even a piece of cardboard. Then, proceed to set up autoradiography (see Chapter 14).

Note: Signal intensity on autoradiograms is most easily quantified through the use of image analysis software, after first ensuring that the observed signal is within the linear range of the film. This can be easily accomplished by varying the exposure time to several pieces of film.

4. Following autoradiography, the hybridized RNA probe can be stripped from the membrane-bound target DNA by washing the filters in 0.4 N NaOH at 42° for 30 min, followed by a 60 min wash in 1× SSC, 0.5% SDS at 65°. This technique removes all hybridized RNA, leaving the efficiency of subsequent hybridization unimpaired.
5. If the filter membrane is not to be used for subsequent hybridization, the magnitude of the observed autoradiographic signal can be precisely quantified by simply cutting the appropriate samples from the filter membrane strip and counting directly in scintillation fluor.

Protocol: alternative procedure for nuclear run-off assay

Several simplified versions of the traditionally cumbersome and labor-intensive nuclear run-off assay has been described (Celano *et al.*, 1989b; Fei and Drake, 1993), a modification of one of which (Celano *et al.*, 1989b) is included here. In this procedure, labeled RNA transcripts are purified rapidly by chaotropic acid–phenol–guanidinium thiocyanate extraction (Chomczynski and Sacchi, 1987; see Chapter 2) and without TCA precipitation. Be sure to wear gloves throughout and observe good RNase-free technique.

1. Isolate nuclei as described above, or elsewhere (Celano *et al.*, 1989a; deBustros *et al.*, 1986).
 2. Collect isolated nuclei (up to 5×10^7) in sterile 2.2 ml microfuge tubes in 160 μ l of nuclear storage buffer (40% glycerol; 50 mM Tris, pH 8.3; 5 mM MgCl₂; 0.1 mM EDTA). Quick-freeze the nuclei in an ethanol-dry ice bath and store at -80° for up to 1 year. *Do not* store this type of tube in liquid nitrogen. Alternative, aliquots of nuclei can be stored in cell freezing ampoules, tightly sealed, and stored in liquid nitrogen for up to a year.
 3. Thaw aliquots of frozen nuclei on ice. For each 200 μ l of nuclei, add 25 μ l of 10 \times run-off buffer (40 mM ATP, GTP, CTP; 5 mM DTT; 50 mM MgCl₂; 800 mM KCl) and 20 μ l [α -³²P]-UTP (200 uCi; 3000 Ci/mmol). Incubate for 15 min at 26 $^\circ$.
- Note: Using an NTP stock solution from any of several vendors will give the greatest reproducibility from one experiment to the next. Because NTP stock solutions prepared in the lab are lacking in quality control, so to speak, the labeling efficiency is likely to be impacted.*
4. Lyse the nuclei and initiate DNA digestion by adding 12 μ l 20 mM CaCl₂ and 12 μ l 10 mg/ml RNase-free DNase I. Incubate for 5 min at 26 $^\circ$.
 5. Add 30 μ l of 10 \times SET (5% SDS; 50 mM EDTA; 10 mM Tris, pH 7.4) and 7 μ l 10 mg/ml carrier RNA (e.g., junk yeast RNA). Initiate peptide hydrolysis with the addition of 3 μ l 10 mg/ml proteinase K. Incubate the samples at 37 $^\circ$ for 30 min.
 6. Shear genomic DNA by drawing the lysate carefully and repeatedly through a 25-gauge needle positioned on a 1 ml tuberculin syringe.
 7. Extract the RNA from the sample by adding the following to each tube containing labeled nuclear RNA (mix after the addition of each):
 - 500 μ l GTC (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 7.0; 0.5% sarcosyl; 0.1 M 2-mercaptoethanol)
 - 80 μ l sodium acetate (2 M, pH 4.8)
 - 800 μ l water-saturated phenol
 - 160 μ l chloroform:isoamyl alcohol (49:1).Incubate samples on ice for at least 15 min.
- Note: If desired, commercially available formulations for RNA isolation, such as Trizol or TRI Reagent[®], can be substituted in this step. Follow the instructions that accompany this product for efficient isolation of RNA from this step forward.*
8. Microcentrifuge samples at 12,000 \times g for 10 min, at 4 $^\circ$ if possible.
 9. Transfer the upper aqueous phase to a fresh microfuge tube, taking care to avoid the interface. Add an equal volume of isopropanol to this aqueous material to precipitate the RNA. Store sample(s) for at least 1 h at -80° or overnight at -20° .
 10. Collect the precipitate RNA by microcentrifugation at 12,000 \times g for 10 min, at 4 $^\circ$ if possible.

Note: It is recommended that the RNA pellet be redissolved in 300 μ l of guanidinium thiocyanate solution (see step 7) and transferred to a 1.7 ml microfuge tube. The RNA is then reprecipitated by the addition of 300 μ l of isopropanol and collected by centrifugation as described above.

11. Wash the pellet two or three times with 70% ethanol (95% ethanol diluted in sterile H₂O) and briefly air-dry. If desired, a final wash with 95% ethanol will accelerate the drying process. In either case, do not allow the RNA to dry completely. Dissolve the RNA pellet in 250–500 μ l in a solution of 10 mM Tris, pH 7.2; 1 mM EDTA; 0.1% SDS.
12. Determine the activity of the RNA probe by counting a 2 μ l aliquot. The total expected activity is approximately 10⁷ cpm for the assay when starting with 5×10^7 cells.
13. Hybridize labeled RNA to DNA target sequences as described above, or elsewhere (Celano *et al.*, 1989a; deBustros *et al.*, 1986).

Protocol: nuclease protection–pulse label transcription assay

The mechanics of the nuclear run-off assay (label incorporation into nascent transcripts in isolated nuclei) differ markedly from pulse labeling techniques. Subjecting cell cultures to a 5-min pulse with [³H]-uridine (Nevins and Darnell, 1978) would require a sufficiently high rate of transcription in order to derive meaningful incorporation of label. This certainly is not the case with single-copy genes or other sequences that are transcribed at basal levels. This issue and other concerns pertaining to the inherent difficulties of preparing nuclei capable of meaningful label incorporation have been partially circumvented by coupling the nuclease protection assay with the pulse-labeling of nuclear RNA (Greene and Pearson, 1994).

This technique involves pulse labeling tissue culture cells for 10 min, followed by RNA isolation, solution hybridization with cold (unlabeled) gene-specific sequences, and finally incubation with S1 nuclease. With this approach, cDNA probes for specific genes produce significantly higher fidelity data than oligo-nucleotide probes. The nuclease-resistant hybrids are precipitated and the relative rate of transcription inferred from the incorporated label, as determined by liquid scintillation counting. As with many transcription assays, this protocol has the advantage of labeling transcripts while maintaining natural nuclear geometry and, in this protocol, maintaining cellular geometry as well. At the conclusion of the labeling period, RNA is isolated from the cells for hybridization analysis. In short, this is a rapid, highly quantitative assay. Moreover, without the requirement for X-ray film, the inherent limitations associated with autoradiography (extended exposure time; non-linearity of the film) are superseded by scintillation counting. Thus, this assay has the ability to detect subtle changes in transcriptional activity that might otherwise be missed by the more traditional nuclear run-off approach.

1. Cells growing in tissue culture are labeled with the addition of 25 μ Ci/ml [³H]-uridine (specific activity 27.1 Ci/mmol) added directly to the growth medium tissue culture vessel.

Note: RNA from 10^7 cells is generally required for one assay in duplicate, including all controls (see step 11). At the conclusion of the labeling period, be sure to dispose of radioactive tissue culture media and RNA isolation by-products in a safe manner, according to departmental guidelines.

2. Harvest cells and wash cell pellets twice with ice-cold PBS. All subsequent centrifugations are carried out at 4°.
3. Resuspend cells (up to 10^7) in 1.8ml NP-40 lysis buffer (10mM Tris, pH 7.4; 10mM NaCl; 3mM MgCl₂; 0.5% NP-40). Transfer the cell suspension to an autoclaved 2.2ml microfuge tube. Incubate on ice for 3 min.
4. Centrifuge at $4000 \times g$ for 5 min at 4° to collect nuclei.
5. Carefully remove and discard supernatant. Resuspend nuclei in 800 μ l SDS buffer (0.5% SDS; 1mM CaCl₂; 5mM MgCl₂; 20mM HEPES, pH 7.5).
6. Add 50U RNase-free DNase I and incubate for 15 min at 37°.
7. Split the nuclear lysate into two microfuge tubes, and then carefully extract each with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1).
8. Centrifuge to separate the phases. Carefully transfer the upper aqueous phase to a fresh tube.
9. Precipitate RNA with the addition of 0.1 volume 3M sodium acetate, pH 5.2, and 2.5 volumes of 95% ethanol. Store at -20° for several hours to overnight.
10. Carefully wash sample once with 70% ethanol (95% ethanol diluted with sterile H₂O) and centrifuge again if necessary. Briefly air dry the sample but do not allow it to dry out.
11. Resuspend the RNA pellet in 125 μ l S1 hybridization buffer (40% deionized formamide; 40mM PIPES, pH 6.4; 400mM NaCl; 1mM EDTA, pH 8.0). Be sure to redissolve the RNA by repeated pipetting. This is a critical step. Do not vortex.

Note 1: Dissolve the RNA pellet in enough hybridization buffer to allow 20 μ l for each hybridization reaction to be performed and run each sample in duplicate.

Note 2: Appropriate controls to be performed include (1) a tube with no probe and no S1 nuclease; (2) a tube with no probe and 250 U S1 nuclease; and (3) a tube with probe, but no S1 nuclease.

12. Heat each resuspended RNA sample to 50° for 2 min.
13. Add denatured cold probe (unlabeled) to the samples to a final concentration of 25ng/ μ l, or 0.5 μ g probe per 20 μ l hybridization reaction.

Note: Be sure to use linearized, denatured DNA. Double-stranded probes can be easily denatured by boiling them for 10min and then plunging them into ice until further use. Alternatively, the RNA/probe mix can be heated to 65° for 10min to ensure all-inclusive denaturation at the onset of the hybridization.

14. Hybridize overnight at an experimentally determined temperature. For optimization suggestions, see Chapter 13. If uncertain where to begin, hybridize at 37°.
15. At the conclusion of the hybridization period, add to each sample:
160 μ l 2 \times S1 buffer (0.5M NaCl; 0.1M sodium acetate, pH 5.5; 9mM ZnSO₄)
142 μ l H₂O
6 μ l boiled salmon DNA (10mg/ml; [Final] \approx 0.02 μ g/ml)
12 μ l S1 Nuclease (25U/ μ l; [Final] \approx 1U/ μ l).

Incubate at 32° for 1 h.

Note: Be sure to prepare one control tube without the addition of any S1 nuclease. This will provide the total counts in the system.

16. Add 80 μ l stop buffer (3M sodium acetate, pH 5.2; 20mM EDTA, pH 8.0; 40 μ g/ml carrier RNA) to each sample.

17. Add 1.0ml ice-cold 95% ethanol to each tube. Invert to mix thoroughly and then incubate on ice for 20 min.
18. Collect precipitate on GF/F glass fiber filters using a Millipore Sampling Manifold (Millipore, Billerica, MA), occasionally referred to as a vacuum filtration bank.
19. Wash each filter twice with 500 μ l cold 70% ethanol.
20. Transfer filters to scintillation vials and add an appropriate volume of fluor. Count samples in triplicate for 5 min. The activity remaining on each filter is directly proportional to the extent of hybridization between probe and target RNA, and hence proportional to the relative rate of synthesis.

Distinguishing among the activities of RNA polymerases

The relative rate of transcription of several genomic sequences can be evaluated simultaneously by the nuclear run-off assay. In eukaryotic cells, the products of transcription are, collectively, the result of four different RNA polymerases. The activities of these enzymes may be distinguished from each other by using the fungal cyclic octapeptide α -amanitin, from the very poisonous mushroom *Amanita phalloides*. This toxin binds to RNA polymerase molecules and is one to which the four eukaryotic RNA polymerases exhibit differential sensitivity (Roeder, 1976). For example, at a concentration of 0.01 μ g/ml, RNA polymerase II activity is reduced by 50% (Roeder, 1976). RNA polymerase III is also sensitive to the inhibitory activity of α -amanitin, although a concentration of 25 μ g/ml is required to achieve an equivalent 50% inhibition of activity. In contrast, RNA polymerase I exhibits no sensitivity to this peptide. A fourth RNA polymerase (single-polypeptide nuclear RNA polymerase IV, or spRNAP-IV) has recently been discovered in mammals (Kravchenko *et al.*, 2005), and RNA polymerase IV (RNAP IV, or RNA pol IV) has likewise been discovered in plants (Onodera *et al.*, 2005). These enzymes are rather insensitive to α -amanitin, as are the RNA-dependent RNA polymerases (RdRPs) that are associated with miRNA expression. While α -amanitin is extremely toxic and must be handled with great care, it is a very useful tool for determining the amount of RNA synthesis attributable to each polymerase. The inclusion of α -amanitin in one or more labeling reactions is an important and an excellent negative control by which blot hybridization signals are validated as authentic and not artifactual.

In nuclei isolated from cultured cells, the contribution to the total observed RNA synthesis due to RNA polymerase I, RNA polymerase II, and RNA polymerase III is about 50–70%, 20–40%, and about 10%, respectively. Clearer definition may be given to any model system by performing three separate labeling reactions in which α -amanitin is added as a reaction buffer component *before* the addition of the nuclei:

1. no α -amanitin added to labeling reaction;
2. 1 μ g/ml α -amanitin added to the labeling reaction;
3. 100 μ g/ml α -amanitin added to the labeling reaction.

Following addition of nuclei, the reaction mixtures are briefly incubated on ice and then transferred to the temperature at which the labeling will take place. The difference in the mass of transcription products in Reactions 1 and 2 represents RNA polymerase II activity. The difference between Reactions 2 and 3 represents RNA polymerase III activity. The activity in Reaction 3 is due to RNA polymerase I activity. These differences associated with α -amanitin poisoning may be assessed by dot-blot analysis (Appendix N), by the separation of transcription products on a denaturing agarose gel, or by electrophoresis on a denaturing 10% polyacrylamide gel (Appendix M). Electrophoresis will permit visualization of the size distribution of all transcription products by autoradiographic exposure directly from the gel.

Alternatively, the size distribution of transcribed species can be determined by sucrose density gradient centrifugation (Marzluff and Huang, 1984). This is an older and much more time consuming procedure than simple gel electrophoresis and requires specialized equipment.

Briefly:

1. Dissolve RNA samples in 0.1% SDS; 1 mM EDTA, pH 7.5; load 0.5 ml of this onto a 17 ml 10 to 40% (w/v) sucrose gradient made up in 0.1 M NaCl, 1 mM EDTA, 0.1% SDS, and 10 mM Tris, pH 7.5. The gradient must be prepared in a tube suitable for ultracentrifugation.
2. Centrifuge at $90,000 \times g$ for 16 h in a suitable rotor. Under the parameters defined here, examination of all RNA species from 4S to 45S will be possible, with the 28S RNA sedimenting 60% of the way down the gradient.
3. Gradients may then be fractionated by UV absorbance at 254 nm, for example with an ISCO UA-6 monitor/fraction collector, and the radioactivity in each fraction may be determined by assaying TCA-precipitable counts in a 10 μ l aliquot from each fraction.

Extraction of nuclear RNA for steady-state analysis

The nuclear run-off assay facilitates the assessment of the transcription rate of individual genes by the elongation of nascent polyribonucleotides in the presence of labeled NTP precursor. In addition, steady-state nuclear RNA, as well as labeled RNA produced by nuclear run-off are essential for understanding the processing of the primary transcript. Northern analysis of nuclear RNA, for example, yields quantitative data based on signal intensity of discrete bands, but also provides a qualitative component that cannot be discerned by PCR or nuclease protection. The appearance of a reproducible banding pattern strongly supports the existence of a primary RNA transcript and, in so doing, attests to the reliability of the nuclear run-off assay. A series of bands on an autoradiogram is associated with the systematic splicing together of coding (exon-associated) sequences and concomitant removal of non-coding (intron-associated) sequences.

In order to assess the overall size distribution of labeled nuclear transcripts, an aliquot of the reaction mixture can be studied by minigel electrophoresis coupled with in-gel autoradiography. Alternatively, target DNA may be

digested, electrophoresed, and Southern blotted onto a filter membrane, followed by attempted hybridization with the heterogeneous population of labeled nuclear RNA (a reverse Northern blot, of sorts). A major drawback of this technique, however, is the lack of information about the size of the nuclear RNA molecules themselves, since it is the DNA molecules that are electrophoresed, rather than the RNA. In extreme cases, partial degradation of full-length hnRNA molecules may well result in non-specific hybridization of the resulting oligoribonucleotides to target DNA molecules with which full-length, undegraded hnRNAs would not normally hybridize. High-stringency washes, in conjunction with RNase digestion of non-specific RNA molecules, will alleviate some, but not all, of the background due to this phenomenon.

The isolation of steady-state nuclear RNA begins with the preparation of nuclei from whole cells or tissue samples as described previously. Invariably the highest quality RNA always results from the most expedient extraction procedures, i.e., those which minimize the interval between cellular disruption and immobilization of RNA on a filter membrane. Inhibition of RNase activity in problematic cells can be accomplished by incorporating guanidinium-based techniques (Chirgwin *et al.*, 1979; Chomczynski and Sacchi, 1987). Once the nuclei have been isolated from the cytosolic contents, one may elect to purify the RNA from the nuclei as if starting with intact cells. The imperative to control nucleases persists for nuclear RNA isolation. Many of the protocols describing the isolation of nuclear RNA recommend heating the sample in the presence of phenol, accompanied by vigorous shaking to shear the genomic material. However, excellent quality RNA can also be recovered from biological sources through the use of silica-based products described in Chapter 2.

Protocol: direct isolation of nuclear RNA

The following protocol is a modification of the procedure of Soeiro and Darnell (1969) in which the use of hot phenol for RNA purification was originally described.

1. Thaw previously frozen nuclei on ice and pellet at $750 \times g$ for 3 min at 4° . If starting with freshly prepared nuclei, proceed to step 2.
2. Wash nuclear pellet with 1 ml ice cold modified RSB/K buffer (10 mM Tris, pH 7.9; 10 mM NaCl; 10 mM $MgCl_2$; 100 mM KCl). Transfer nuclei to an RNase-free polypropylene tube.

Note: If this extraction is scaled down, it is possible to isolate the RNA in a 2.2 ml microcentrifuge tube. As always, it is best to keep nucleic acids concentrated, working in the smallest possible volumes.

3. Resuspend the nuclear pellet to a concentration of $1-5 \times 10^7$ nuclei/ml in HSB buffer (10 mM Tris, pH 7.4; 500 mM NaCl; 50 mM $MgCl_2$; 2 mM $CaCl_2$), to which RNase-free DNase I has been added to a final concentration of 50 U/ml.

Note: Add RNase-free DNase I just prior to use.

4. Pipette this mixture up and down for 30 s at room temperature. This is necessary for lysate homogeneity.

Note: The objective of this brief treatment with DNase I is to partially fragment, rather than completely degrade, the genomic DNA. The decrease in viscosity increases the manageability of the prep. The DNA will be selectively removed later in this procedure.

5. Add an equal volume of SDS extraction buffer (10 mM Tris, pH 7.4; 20 mM EDTA; 1% SDS).
6. Add an equal volume of phenol saturated with NETS buffer (10 mM Tris, pH 7.4; 100 mM NaCl; 10 mM EDTA; 0.2% SDS).

Note: See Appendix D for tips on phenol saturation.

7. Add an equal volume of chloroform or a mixture of chloroform:isoamyl alcohol (24:1). Mix thoroughly and *carefully*.
8. Heat to 55° for 10 min, with periodic shaking.
9. Cool on ice or in an ice-water bath for 5 min. Centrifuge at $2500 \times g$ for 3 min to separate the phases.
10. Remove the lower organic phase by carefully sliding a sterile silanized Pasteur pipette or microcentrifuge tip along the side of the tube, through the aqueous phase and interphase, and aspirate from the bottom of the tube. Leave the aqueous phase and interphase in the tube and repeat steps 6 and 7.

Note 1: This approach is especially important when working with small quantities of nuclear RNA, where there is a potential for loss of RNA by entrapment in chromatin.

Note 2: See Appendix I for instructions on silanizing glassware.

11. Centrifuge at $2500 \times g$ for 3 min to separate phases.
12. Carefully recover the aqueous phase, taking care not to disrupt the lower organic phase or interphase, if present. Add an equal volume of chloroform to the aqueous material, mixing thoroughly and carefully. Pulse centrifuge briefly to separate the phases.
13. Transfer the aqueous phase to a new tube. Precipitate the nuclear RNA at -20° with the addition of 2.5 volumes of ice-cold 95% ethanol.
14. Collect the precipitate by centrifugation at 4°. Wash at least once with 70% ethanol (95% ethanol diluted with sterile H₂O) to remove excess salt. Quantify yield and perform quality control testing as described in Chapter 6.

Protocol: preparation of nuclear RNA from cells enriched in ribonuclease

The following is a modification of the procedure of Chomczynski and Sacchi (1987). This method is the basis of RNA isolation products TRI Reagent and TRIzol. For the fainthearted who prefer not to prepare individual reagents in-house, either of the above-referenced products can be substituted with confidence in place of the following protocol.

1. Prepare nuclei as described above; nuclei can be prepared ahead of time and stored frozen. Allow frozen nuclei to thaw on ice for 3–5 min.
2. Lyse intact nuclei with the addition of GTC solution (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 7.0; 0.5% sarcosyl; 100 mM 2-mercaptoethanol). Use 100 μ l per 2×10^6 nuclei.

Note: The volumes indicated for the remaining steps in this procedure assume a starting volume of 1 ml of GTC solution D. For scaled-up extractions, remember to increase the volumes of all other reagents appropriately. If less than 800 μ l of GTC solution is used, the entire extraction can be carried out in a 2.2 ml microfuge tube.

3. Transfer the homogenate to a disposable **polypropylene tube** (15 ml size works the best) and add, in order:
 - 0.1 ml 2 M sodium acetate, pH 5.2
 - 1.0 ml water-saturated phenol (molecular biology quality)
 - 0.2 ml chloroform:isoamyl alcohol (49:1).

Cap tube and mix *carefully* and thoroughly by inversion following the addition of reagent. Shake tube vigorously for 10 s after all reagents have been added.

4. Cool sample on ice for at least 15 min and then centrifuge at 4° to separate the phases.
Note: Be sure to check the manufacturer's maximum g-force rating before use.
5. Transfer aqueous phase (contains RNA) to a fresh tube and mix with 0.75 volume of isopropanol (approximately 1 ml). Store at -20° for at least 1 h to precipitate RNA.

Note: The RNA should be precipitated in a tube that can withstand the required 10,000 \times g centrifugation in the next step. If the volume is too great to fit into one or a few microcentrifuge tubes, then acid-washed Correx glass tubes, used with the proper rubber adapters, are a suitable alternative.

6. Collect precipitate by centrifugation at 10,000 \times g for 20 min at 4°. Carefully decant and discard supernatant.
7. Completely dissolve RNA pellet in 300 μ l of GTC solution (see step 2) and then transfer to an RNase-free 1.5 ml microcentrifuge tube.
8. Reprecipitate the RNA by adding 0.75 volume of ice cold isopropanol at -20° for 1 h.
9. Collect precipitate at top speed in a microcentrifuge for 10 min at 4°. Carefully decant and discard supernatant.
10. Wash pellet with 70% ethanol and centrifuge again. The pellet may be washed once more with 95% ethanol to accelerate air-drying. Do not allow the sample to dry out completely.
11. Redissolve damp RNA pellet in 50 μ l of TE buffer, pH 7.5 or 50 μ l of sterile H₂O. Incubation at 65° for 10 min may facilitate solubilization. Determine sample concentration (Chapter 6); store RNA in suitable aliquots at -80°. As always, avoid repeated freezing and thawing.

Troubleshooting nuclear RNA analysis

If difficulties are encountered when performing nuclear RNA analysis, consider the following:

1. Poor label incorporation may be due to incomplete cell lysis. Cells must be lysed in order for the labeled nucleotide(s) to gain access to the nucleus. Use a drop of trypan blue to assess lysis. The nuclei will look like blue ovals under a microscope.
2. Poor label incorporation may also be due to the incubation temperature. Because transcription in isolated nuclei appears to be shut down after a brief period at

- 37° (Marzluff *et al.*, 1973), most investigators perform the labeling step between room temperature and 30°.
3. If the assay does not appear to be linear, check to ensure the cold DNA probes immobilized on the filter(s) were in a molar excess relative to the labeled transcripts. That the target DNA is in excess can be determined by (a) varying the amount of input RNA, in which case the hybridization signal should be proportional to the amount of input RNA; (b) varying the amount of target DNA on the filter paper, in which case there should be no effect on the amount of RNA hybridized if the DNA is in excess; and (c) rehybridizing RNA that does not bind to the target DNA or the filter paper, none of which should participate in hybridization.
 4. To further investigate the linearity of the assay, make triplicate filter membrane strips, onto which samples of target DNA have been immobilized. Prepare multiple two-fold dilutions of the heterogeneous labeled run-off RNA (based on measured specific activity) and then hybridize to individual filter paper strips containing the blotted DNAs. The signal intensity obtained upon hybridization detection should, of course, reflect the twofold dilutions of the labeled RNA.

References

- Benecke, B.J., A. Ben-Ze'ev and S. Penman. (1978). The control of mRNA production, translation, and turnover in suspended and reattached anchorage-dependent fibroblasts. *Cell* **14**, 931-939.
- Ben-Ze'ev, A., S.R. Farmer and S. Penman. (1980). Protein synthesis requires cell-surface contact while nuclear events respond to cell shape in anchorage-dependent fibroblasts. *Cell* **21**, 365-372.
- Celano, P., S.B. Baylin and R.A. Casero. (1989a). Polyamines differentially modulate the transcription of growth-associated genes in human colon carcinoma cells. *J. Biol. Chem.* **264**, 8922-8927.
- Celano, P., C. Berchtold and R.A. Casero. (1989b). A simplification of the nuclear run-off assay. *BioTechniques* **7**(9), 942-945.
- Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald and W.J. Rutter. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.
- Chomczynski, P. and N. Sacchi. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.* **162**, 156-159.
- deBustros, A., S.B. Baylin, M.A. Levin and B.D. Nelkin. (1986). Cyclic AMP and phorbol esters separately induced growth inhibition, calcitonin secretion, and calcitonin gene transcription in cultured human medullary thyroid carcinoma. *J. Biol. Chem.* **261**, 8036-8041.
- Derman, E., K. Krauter, L. Walling, C. Weinberger, M. Ray and J.E. Darnell. (1981). Transcriptional control in the production of liver-specific mRNAs. *Cell* **23**, 731-739.
- Ernest, M.J., G. Schultz and P. Feigelsen. (1976). RNA synthesis in isolated hen oviduct nuclei. *Biochemistry* **15**, 824-829.
- Farrell, Jr., R.E. and J.J. Greene. (1992). Regulation of *c-myc* and *c-Ha-ras* oncogene expression by cell shape. *J. Cellular Physiol.* **153**, 429-435.

- Fei, H. and T.A. Drake. (1993). A rapid nuclear run-off transcription assay. *BioTechniques* 15, 838.
- Greenberg, M.E. (1988). Identification of newly transcribed RNA. In *Current Protocols in Molecular Biology* (F.M. Ausubel, R. Brent, R.E. Kingston, *et al.*, Eds). John Wiley & Sons, New York. p. 4.10.1.
- Greenberg, M.E. and E.B. Ziff. (1984). Stimulation of 3T3 cells induces transcription of the *c-fos* proto-oncogene. *Nature* 311, 433–438.
- Greene, J.J. and S.L. Pearson. (1994). Measurement of gene-specific transcription by nuclease protection of pulse-labeled nuclear RNA. *J. Biochem. Biophys. Methods* 29, 179–187.
- Groudine, M., M. Peretz and H. Weintraub. (1981). Transcriptional regulation of hemoglobin switching on chicken embryos. *Mol. Cell. Biol.* 1, 281–288.
- Gurney, T. and D.N. Foster. (1977). Nonaqueous isolation of nuclei from cultured cells. *Methods Cell Biol.* 16, 45–68.
- Jelinek, W., G. Molloy, R. Fernandez-Munoz, M. Salditt and J.E. Darnell. (1974). Secondary structure in heterogenous nuclear RNA: Involvement of regions from repeated DNA sites. *Mol. Biol.* 82, 361–370.
- Kanazawa, A., M. O'Dell, R.P. Hellens, E. Hitchin and M. Metzloff. (2000). Mini-scale method for nuclear run-on transcription in plants. *Plant Mol. Biol. Rep.* 18, 377–383.
- Kwei, K.A. (2004). Transcriptional repression of catalase in mouse skin tumor progression. *Neoplasia* 6, 440–448.
- Lund, E. and P.L. Paine. (1990). Nonaqueous isolation of transcriptionally active nuclei from *Xenopus* oocytes. *Methods Enzymol.* 181, 36–43.
- Marzluff, W.F. (1973). Transcription of ribonucleic acid in isolated mouse myeloma nuclei. *Biochemistry* 12, 3440–3446.
- Marzluff, W.F. (1978). Transcription of RNA in isolated nuclei. *Methods Cell Biol.* 19, 317–332.
- Marzluff, W.F. and R.C.C. Huang. (1975). Chromatin directed transcription of 5S and tRNA genes. *Proc. Natl. Acad. Sci.* 72, 1082–1086.
- Marzluff, W.F. and R.C.C. Huang. (1984). Transcription of RNA in isolated nuclei. In *Transcription and Translation: A Practical Approach* (B.D. Hames and S.J. Higgins, Eds). IRL Press, Oxford. p. 89.
- Marzluff, W.F., E.C. Murphy and R.C.C. Huang. (1973). Transcription of ribonucleic acid in isolated mouse myeloma nuclei. *Biochemistry* 12, 3440–3446.
- Marzluff, W.F., E.C. Murphy and R.C.C. Huang. (1974). Transcription of the genes for 5S ribosomal RNA and transfer RNA in isolated mouse myeloma cell nuclei. *Biochemistry* 13, 3689–3696.
- Nevins, J.R. (1987). Isolation and analysis of nuclear RNA. *Methods Enzymol.* 152, 234–241.
- Nevins, J.R. and J.E. Darnell. (1978). Steps in the processing of Ad2 mRNA: Poly(A)⁺ nuclear sequences are conserved and poly(A)⁺ addition precedes splicing. *Cell* 15, 1477–1493.
- Ng, D.C., M-J. Su, R. Kim and D.D. Bikle. (1996). Regulation of involucrin gene expression by calcium in normal human keratinocytes. *Frontiers Biosci.* 1, 16–24.
- Patrone, G., F. Puppo, I. Ceccherine, A. Puliti and R. Ravazzolo. (2000). Nuclear run-on assay using biotin labeling, magnetic bead capture and analysis by fluorescence-based RT-PCR. *BioTechniques* 29, 1012–1017.

- Price, R. and S. Penman. (1972). A distinct RNA polymerase activity, synthesizing 5.5S, 5S, and 4S RNA in nuclei from adenovirus 2-infected HeLa cells. *J. Mol. Biol.* **70**, 435–450.
- Roeder, R.G. (1976). Nuclear RNA polymerase. In *RNA Polymerase* (R. Losick and M. Chamberlain, Eds). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. pp. 283–329.
- Soeiro, R. and J.E. Darnell. (1969). Competition hybridization by “pre-saturation” of HeLa cell DNA. *J. Mol. Biol.* **44**, 551.
- Wittelsberger, S.C., K. Kleene and S. Penman. (1981). Progressive loss of shape-responsive metabolic controls in cells with increasingly transformed phenotype. *Cell* **24**, 859–866.

17 cDNA Synthesis

Rationale	367	Protocol: first-strand cDNA synthesis	377
cDNA synthesis – an overview	368	Assessing cDNA synthesis efficiency	379
First-strand considerations	369	Cloning cDNA	379
Reverse transcriptase options	372	Ligation considerations	381
Second-strand considerations	375	Enzymes used for ligation	383
Classical methods	376	Applications	383
PCR-based methods	377	References	384

Rationale

RNA is an easily accessible molecule within the cell and therefore a commonly assayed parameter of gene expression. The numerous strategies for cell- and tissue disruption in the preceding chapters can be utilized confidently to support the expedient isolation of RNA from various biological sources. If adequate attention is given to the subtleties of isolation and post-isolation handling, the RNA will ordinarily be of sufficient quality to support any of a number of assays for the evaluation of the transcriptional activity of one or several genes.

The classical methods of Northern analysis, nuclease protection, and nuclear runoff all measure RNA abundance directly through the use of DNA and RNA probes. It must be emphasized once again, however, that there is an upper limit to the sensitivity that each of these methods can offer. This ceiling is due to two intrinsic difficulties associated with the RNA, specifically (1) the RNA is naturally labile; and (2) while variable in abundance, individual RNA molecules cannot be amplified directly in order to boost the signal. Happily, there are straightforward solutions to these difficulties. First, RNA can be converted enzymatically into complementary DNA (cDNA) with minimal fanfare. Second, the technology now exists to greatly amplify cDNA (and genomic DNA) with amazing speed and robustness. That cDNA is much more stable than the corresponding RNA template material from which it is derived facilitates extensive analysis in both the short- and long-term by any of a number of techniques. The golden rule of cDNA synthesis is this: when you make cDNA, what you are really doing is making a permanent biochemical record of the cell. One makes a biochemical “snapshot” since the only cDNA that *can* be synthesized comes from the RNA of transcriptionally active loci.

cDNA synthesis – an overview

The synthesis of cDNA is a central component of research involving molecular biology techniques. cDNA is useful in both the single-stranded- and double-stranded forms, and is chemically equivalent to genomic DNA, though there are often structural differences. While the classical methods for cDNA propagation have given way to PCR, an historical perspective of the mechanics of traditional cDNA synthesis and its significance remains central to the proper understanding of the technology, its flexibility, and the interpretation of data.

cDNA is a product of enzymatic *in vitro* synthesis, one strand at a time, using RNA as template material. The need for high quality RNA is the first and foremost requirement. Generally less than a microgram of RNA is needed for many routine applications, and some of the newer techniques require less than 100 ng of starting material. Although the selection of poly(A)⁺ RNA as the starting template for the synthesis of cDNA has historical significance, it is now commonplace to synthesize cDNA directly from total cellular RNA or total cytoplasmic RNA without prior poly(A)⁺ selection. Given the power of PCR, poly(A)⁺ selection is often neither necessary nor recommended because the mechanics of poly(A)⁺ selection often result in the loss and further under-representation of very low abundance transcripts. Thus, methods such as this that are intended to enrich in favor of poly(A)⁺ mRNA are often counterproductive. In any event, the selection of poly(A)⁺ mRNA from previously purified cellular or cytoplasmic RNA may be performed, if desired.

The synthesis of cDNA offers many advantages to the investigator wishing to characterize gene structure or expression, develop nucleic acid probes, or express proteins for any of a number of reasons. First, RNA is a naturally labile single-stranded molecule, and its conversion into more stable single- or double-stranded DNA facilitates long-term storage of these sequences. Second, by cloning newly synthesized cDNA, the investigator creates a method for propagating the cDNA as needed. This approach is facilitated by the large variety of vector molecules compatible with an equally impressive variety of hosts. Succinctly, the joining or ligation of a collection of cDNAs into a suitable vector, and introduction of that vector into an appropriate host, constitutes a cDNA library, which is an archive of sorts that contains a record of all of the gene sequences associated with a particular sample. Third, cDNA molecules, both long and short, can be used to screen (hybridize to) members of much more complex genomic DNA libraries. This approach facilitates the isolation of exon and intron sequences from the structural portion of genes as well as the flanking 5' and 3' sequences. Fourth, since cDNA cannot be synthesized from mRNA that is absent, the emphasis is on active gene expression. In sharp contrast, libraries made from genomic DNA afford the opportunity to clone all cellular sequences regardless of transcriptional profile.

cDNA is synthesized directly from an RNA template; therefore cDNAs present at the end of a synthesis reaction directly mirror the RNA complexity

at the beginning of the synthesis reaction. Because all nucleated cells within an organism have essentially the same genomic material, nearly identical data are expected, for example, from analysis of a genomic splenocyte library and a genomic hepatocyte library from the same organism. In contrast, both the phenotype and physiology of a particular cell is due to the induction and repression of specific genes and gene relays. While it is certainly reasonable to expect that the cDNA populations from two cell- or tissue types will share at least those sequences necessary for cellular viability, the actual members of a cDNA library are quite unique to the biological source. Moreover, because the RNA makeup of a cell is subject to change upon experimental challenge or even simply by progression through the cell cycle, one should expect that the cDNA from a particular source is also subject to change. The take home lesson: in a given organism the cDNA made from various diploid tissues differs, while genomic DNA (mutations and telomeres notwithstanding), does not.

cDNA is synthesized in a stepwise fashion, always demanding the highest quality RNA possible. A comprehensive listing of the enzymes with a past or current association with cDNA synthesis is Table 17.1. These days, most cDNA synthesis kits are quite good and are widely available; it is highly advisable to invest in a cDNA synthesis kit when performing this technique for the first time because of the sensitive nature of cDNA synthesis reactions. A disproportionate number of man-hours and other laboratory resources are likely to be wasted trying to optimize several reagents, with no success guaranteed. On a per reaction cost basis, these kits are easily the most effective and economical way to make cDNA, and these systems make a lot of cDNA. To enhance the understanding of the reader, the following is a description of the steps involved in the synthesis of cDNA. For the hearty scientist, a cDNA synthesis protocol follows.

First-strand considerations

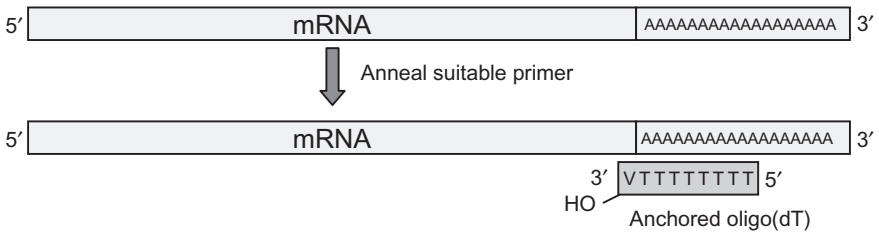
There are two absolute requirements that must be fulfilled in order to support the synthesis of nucleic acids *in vivo* and *in vitro*. First, there must be template (RNA in this case) that can direct the *de novo* synthesis of a complementary strand. Second, there must be a primer molecule that provides a free 3'-OH, a chemical group necessary for all DNA- and RNA polymerases to join together individual nucleotides and support elongation. Keeping in mind that polynucleotides are assembled 5'→3' and that whenever double-stranded molecules form the individual strands must be base-paired in an antiparallel manner, it is clear that the primer molecule providing the necessary 3'-OH must base-pair close to the 3' end of the mRNA template if the synthesis of a full-length cDNA is to be supported. Following heat denaturation to disrupt any RNA secondary structures, a short poly(dT) primer (oligodeoxythymidylic acid), commonly referred to as oligo(dT), is annealed to the poly(A) tail of the RNA. Because of the antiparallel base pairing that will occur between the primer and the RNA template, the primer will provide the 3'-OH group required to support 5'→3'

Table 17.1 The Enzymes of cDNA Synthesis and Cloning*

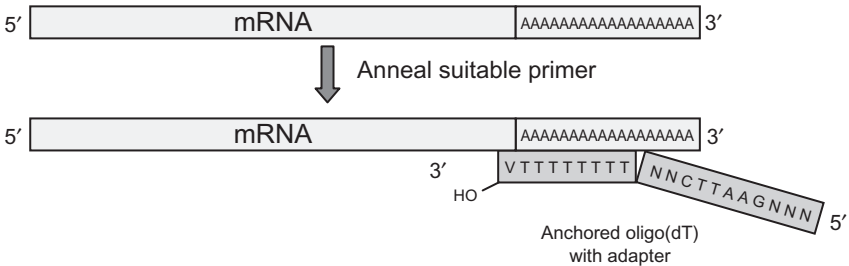
Enzyme	Activities	Role in cDNA synthesis
Reverse transcriptase (RNA-dependent DNA polymerase; RT)	5'→3' polymerase RNase H (minor)	Enzymatic conversion of RNA into cDNA. Substrates may include mRNA, hnRNA, tRNA, rRNA, viral RNA.
DNA polymerase I (Kornberg enzyme)	5'→3' polymerase 5'→3' exonuclease 3'→5' exonuclease	Synthesis of second-strand cDNA by the Gubler and Hoffman (1983) method (5'→3' exonuclease activity removes nicked mRNA).
DNA polymerase I (Klenow fragment)	5'→3' polymerase 3'→5' exonuclease	Synthesis of second-strand cDNA by the classical method of cDNA synthesis.
RNase H	Nicks mRNA component of RNA:DNA hybrid molecules	Nicking activity provides numerous primers (3'-OH) for synthesis of second-strand cDNA.
S1 nuclease	Single-strand-specific nuclease (RNA and DNA)	Removal of hairpins in classical synthesis (enzyme has specificity for single-stranded domains).
T4 DNA ligase	Ligation of cohesive ends; ligation of blunt ends	Seal phosphodiester backbone (...3'-OH + p5'X...).
T4 DNA polymerase	5'→3' polymerase 3'→5' exonuclease	Preparation of blunt ends on ds cDNA; exonuclease activity is 200× greater than DNA polymerase I.
DNA methylase (<i>Eco RI</i> methylase)	Methylase	Methylation of internal restriction sites on cDNA; protects cDNA from cleavage by <i>EcoRI</i> .
Terminal deoxynucleotidyl transferase (Terminal transferase; T.T ⁺ ase)	5'→3' polymerase	Addition of nucleotides to blunt ended (Co ⁺⁺ -dependent) or protruding 3'-OH groups (Mg ⁺⁺ -dependent).
<i>Thermus aquaticus</i> polymerase (<i>Taq</i> polymerase)	Reverse transcriptase 5'→3' polymerase	Poor reverse transcriptase activity for first-strand synthesis; second-strand synthesis/amplification by PCR.
<i>Thermus thermophilus</i> polymerase (<i>Th</i> polymerase)	Reverse transcriptase 5'→3' polymerase	High-temperature first-strand synthesis (Mn ²⁺ -dependent); second-strand synthesis/amplification by PCR (Mg ²⁺).
Other thermostable polymerases	Variable	First- and second-strand synthesis; amplification by PCR.

*The enzymes listed here are the classical enzymes that have been used to support cDNA synthesis, the use of many of which remains commonplace. Several other novel reverse transcriptases have been isolated that support one-tube RT-PCR, some of which are discussed in Chapter 18.

synthesis of a molecule complementary to the RNA template, hereafter known as first-strand cDNA.



Oligo(dT) primers may also be appended at their 5' end with overhanging sequences of all persuasions that will support the addition of restriction enzyme sites or other useful sequences for directional cloning¹. These primers are often referred to as adapters.

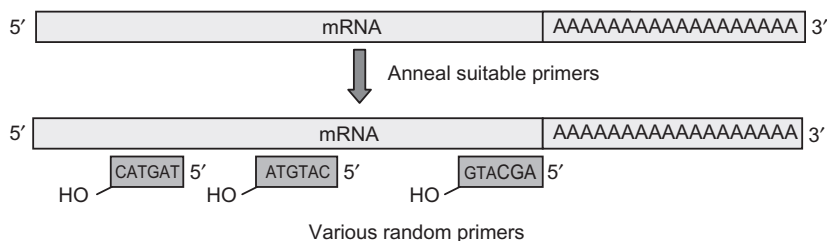


Either of these priming methods fulfills the template and primer requirements needed to support the synthesis of first-strand cDNA. A major improvement in oligo(dT)-priming was the innovative use of an equimolar mixture anchored oligo(dT) primers, having the generic structure $T_{(n)}V$, where $V = A, C, \text{ or } G$. Thus, the sequences of the poly(dT) 13-mer anchored primers are 5' TTTTTTTTTTTTA, 5' TTTTTTTTTTTTC, and 5' TTTTTTTTTTTTG. The anchor nucleotide at the 3' end of the primer (V) forces the primer to base-pair to the 5'-most end of the poly(A) tail. A happy consequence of priming in this manner is the synthesis of much more representative cDNAs, containing greater amounts of the coding portion of each mRNA template molecule. These primers work wonderfully.

It is important to realize from the onset that, if primed as described, the resulting cDNA is really poly(A)⁺-generated cDNA, and not total cDNA, because not all mRNAs are adenylated. cDNA libraries prepared in this fashion do not contain clones corresponding to poly(A)⁻ mRNA. Scientists should be aware that such libraries persist; while no longer common, these libraries

¹Directional cloning refers to the forced ligation of an insert into a vector in one particular orientation rather than the other. This type of cloning is necessary when the investigator plans to express (DNA → RNA → protein) a cloned sequence.

surface periodically when a -80° freezer is (finally) cleaned out and reorganized. When screening a library that may have been synthesized some years ago, knowledge of the method of cDNA priming may influence interpretation of the results. To circumvent the potential shortcomings of oligo(dT)-primed cDNA and cDNA libraries, an aliquot of random hexamers or random nonomers can be used to prime the synthesis of the first strand cDNA. Inclusion of these oligomers results in the internal priming of first-strand cDNA instead of priming only from the 3' end of the template. In general, random-primed cDNA is usually longer and more representative than cDNA primed by oligo(dT) alone.



In this laboratory, the most efficient cDNA synthesis reactions occur when a combination of anchored oligo(dT) *and* random primers are used together in the same reaction tube. This approach eliminates the worry about the possible exclusion of poly(A)⁻ transcripts, and tends to produce a population of much larger cDNA molecules, many of which are close to full length. The parameters governing the synthesis of the first strand need not be changed; the only difference is the simultaneous use of one-half of the recommended mass of *each* type of primer. This is now our standard operating procedure and has been especially valuable in the assay of low abundance genes (Bassett *et al.*, 2000; Bassett *et al.*, 2004). If using total RNA, some investigators prefer to use oligo(dT) alone in order to prevent reverse transcription of any non-mRNAs into cDNA. It is the opinion of this Author that oligo(dT) and random primers together perform in a very satisfactory manner. If the cDNA is intended for PCR, or any other downstream application, and the primers are designed correctly, then using total RNA and double priming is a non-issue.

Random priming also precludes the possibility that cDNA synthesis is primed from the 3'-most region of the poly(A) tail, which can be as long as 200–250 nucleotides: in such an event, a significant portion of the 5' end of the first-strand cDNA would consist of a lot of Ts. Random primers also work very well with mRNAs that exhibit strong secondary structure at the 5' terminus. The important concept here is that failure to produce full-length cDNA molecules results in significant underrepresentation of the 5' end of the transcript.

Reverse transcriptase options

The type of enzyme required for the synthesis of first-strand cDNA is properly known as an “RNA-dependent DNA polymerase”; it is far more

commonly known as reverse transcriptase or simply “RT”. The choice of which reverse transcriptase to use has become a bit more complicated in the past few years and requires some forethought. There are many varieties of reverse transcriptase which are sold under various trade names and in numerous buffer formulations, including AMV, from avian myeloblastosis virus; MMLV, from Moloney murine leukemia virus; HIV RT, from human immunodeficiency virus; various RNase H⁻ reverse transcriptase enzymes; r*Tth*, from *Thermus thermophilus* (Mn²⁺-dependent); and other genetically modified thermostable reverse transcriptases.

The classical enzymes are the AMV and MMLV reverse transcriptases. Most investigators prefer the AMV reverse transcriptase because it is a very processive enzyme and is active at temperatures of 50° or more. It is also a preferred enzyme for reverse transcribing shorter templates, particularly those with more secondary structure. In contrast, MMLV is the favored enzyme when attempting to reverse transcribe longer templates. Unfortunately, some preparations AMV and MMLV reverse transcriptase exhibit high levels of endogenous RNase H activity. The action of RNase H is to nick the RNA component of an RNA:DNA hybrid, precisely that which forms when a DNA primer anneals to template RNA just prior to reverse transcription. The net result is cleavage of the template and concomitant instability of the primer, greatly diminishing the yield of cDNA: the primer cannot be extended because it is no longer associated with an intact mRNA. If reverse transcription is initiated, RNase H activity is sure to interfere with the production of full-length cDNA since the growing hybrid molecule is also a potential substrate. In contrast, MMLV reverse transcriptase generally has lower levels of RNase H activity, and is therefore preferred for reverse transcribing longer templates. MMLV-RT reactions are ordinarily performed between 37° to 40°. In view of this intrinsic difficulty, several genetically engineered reverse transcriptases are available which are certified as being RNase H⁻. These are highly desirable and result in greatly improved first-strand synthesis. The more recently available HIV reverse transcriptase, which has greater thermostability than either the AMV or MMLV enzyme at 50°, is likewise used in situations where cDNA synthesis is somewhat recalcitrant.

Each enzyme has an optimal set of reaction conditions (ionic strength, pH, monovalent cation, divalent cation), all of which are provided by the 10× or 5× first-strand buffer that is specific to the reverse transcriptase being utilized, e.g., AMV or MMLV. This buffer is diluted to a working concentration of 1 × , and the synthesis reaction is conducted at a temperature compatible with the specific reverse transcriptase being used. The investigator must further supplement the reaction with the primer(s), a ribonuclease inhibitor to protect the RNA until it is reverse transcribed, and a nucleotide cocktail consisting of an equimolar mixture of dATP, dCTP, dGTP, and dTTP, hereafter referred to as dNTP. First-strand cDNA synthesis requires anywhere from 10 minutes to 1 hour. Virtually all of the first-strand reaction components currently in use are completely compatible with the chemistries that are subsequently required for amplification of the resulting cDNAs by PCR.

If reverse transcription remains something of a challenge in your lab, here are seven ideas that should be considered for possible inclusion or modification of one's current approach.

1. The integrity of the RNA should be checked. One should always have evidence that the RNA being used in a cDNA synthesis reaction is intact. An elaborate electrophoretic setup is not necessary; all that is required is a minigel separation followed by staining to show the presence of the rRNAs and the degree of smearing within each lane (Chapter 6). This is a simple, but valuable experiment, especially if previously purified RNA is being used after long-term storage.
2. Be sure that the reverse transcriptase formulation being used has been certified as being RNase H⁻ to enhance cDNA synthesis. Examples of these enzymes include the well known Superscript[®] III (Invitrogen), PowerScript[™] (BD Biosciences), and ArrayScript[™] (Ambion). Other vendors offer similar products as well.
3. Include gelatin in the first-strand reaction mix. The function of gelatin, a component in some cDNA synthesis kits, is to stabilize the reverse transcriptase and improve the efficiency of the reaction. Gelatin, when used, is added to a final concentration of 10 µg/ml.
4. Use an enhanced enzyme that is able to reverse transcribe at temperatures greater than 50°. A number of remarkable enzymes are now readily available that can synthesize cDNA at temperatures unheard of just a few years ago; for example, eAMV[™] (Sigma-Aldrich) can be used at 65°. Not only do the higher temperatures promote higher fidelity synthesis of cDNA, they are also very effective at disrupting secondary structures in the RNA, well known for their role in reducing the length of first-strand cDNA.
5. Consider adopting the popular one-tube, one-reaction buffer format, if the cDNA will be used for RT-PCR (Chapter 18). In conjunction with innovative, optimized single-buffering systems, the one-tube method can be exploited such that the entire process of making cDNA and then amplifying it by PCR will require the preparation of one reaction tube (Aatsinki, *et al.*, 1994; Murakawa *et al.*, 1988; Sellner and Turbett, 1998). After sealing, the tube need not be opened again until after the cDNA has been synthesized, amplified, and is ready for analysis.
6. Always prepare a cDNA master mix when multiple cDNA synthesis reactions are to be performed. This is very much the same strategy as in the preparation of a PCR master mix, in which all but one component are pipetted together in a large volume, and then aliquoted into the individual reaction tubes². This approach precludes procedural errors due to pipetting. The reverse transcriptase and RNasin (RNase inhibitor, discussed in Chapter 7) are supplied in a glycerol storage buffer and should never be subjected to repeated freeze/thaw cycles. The other reagents are quite stable at -20° for extended periods.
7. Consider the template-switching approach to cDNA synthesis (Chenckik *et al.*, 1998; Matz *et al.*, 1999). Succinctly, it is the habit of MMLV reverse transcriptase to add a few extra nucleotides (most often up to three deoxycytidines) to the 5' end of first-strand cDNA. This occurs in a non-template-dependent manner and occurs efficiently only if the reverse transcriptase reaches the 5' end of the RNA template.

²For large-scale investigations in this laboratory, 1–2 ml of cDNA master mix is prepared (excluding the reverse transcriptase and RNasin) and stored at -20°. Using this strategy, the investigator has access to identical chemistries for follow-up and repeat applications, thereby minimizing at least one potential source of procedural error.

Consequently, a template-switching oligonucleotide (rG), i.e., an oligo with a guanosine tract at its 3' end, base-pairs with the newly synthesized first-strand poly(dC) tract, causing the reverse transcriptase to switch templates; the primer with the oligo(G) sequence at its 3' end is now the template for continued cDNA synthesis activity. Thus, known sequences are thereby incorporated at the 3' end of first-strand cDNA (Fig 17.1). This methodology strongly enriches for full-length cDNAs for subsequent PCR-based assays. This technology is widely known as SMART™ cDNA synthesis (Clontech) and is a patented process. Contact Clontech for details.

Second-strand considerations

While that the mechanics of first-strand cDNA synthesis are nearly identical to the method used at the inception of the technology, the greatest advances in cloning cDNA pertain to the strategies and innovations associated with the synthesis of the second-strand cDNA. Very early methods involved alkaline hydrolysis of the original mRNA template after synthesis of the first-strand, and actually relied on the transient formation of a hairpin at the 3' end of the first-strand, which, acting as primer, supported the synthesis of the second-strand cDNA by using the first-strand as a template. The resulting double-stranded molecule was joined

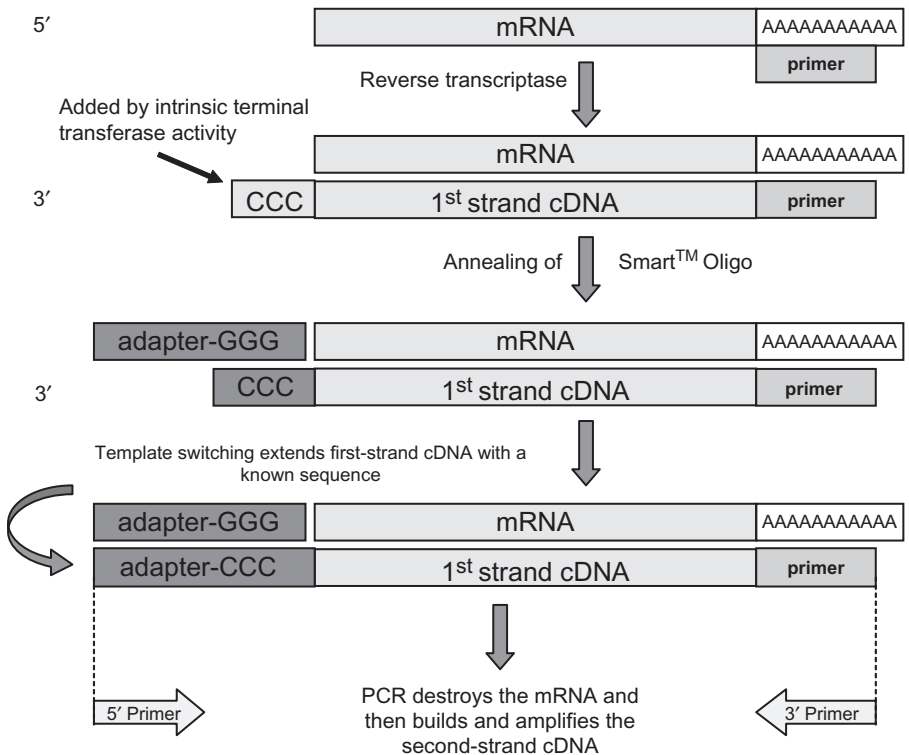
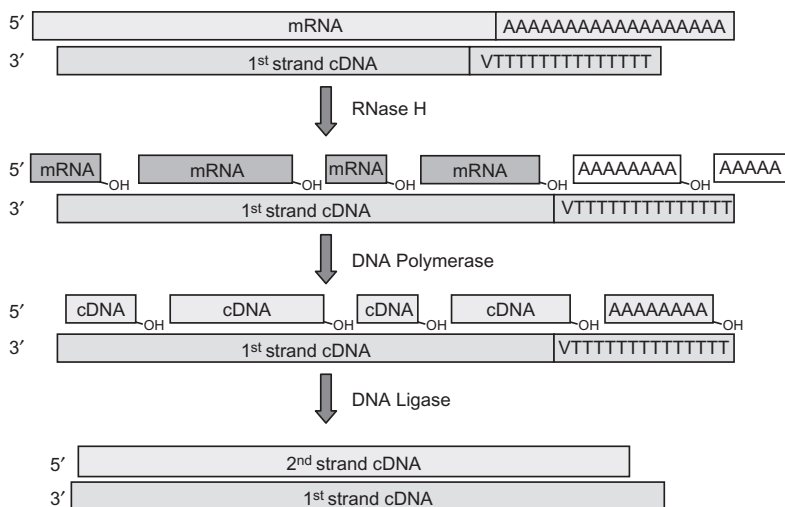


Figure 17.1 Template-switching cDNA synthesis using SMART™ (Clontech) technology.

by a single-stranded hairpin, requiring an S1 nuclease digestion (Chapter 15) to remove the hairpin for subsequent vector insertion (Efstratiadis, *et al.*, 1976; Rougeon and Mach, 1976). This was a primitive approach by today's standards, resulting in the loss of sequences corresponding to the 5' end of the original mRNA, as well as exposing the newly synthesized cDNA to the extremely aggressive and often uncontrollable degradative activity commonly associated with the S1 enzyme. No one does this any more.

Classical methods

Okayama and Berg (1982) subsequently developed an elegant, novel method by which an oligo(dT)-tailed vector was used to prime the synthesis of the first-strand cDNA. The resulting RNA:cDNA was circularized by linker ligation. The real innovation in this technique involved the use of a mixture of the enzymes RNase H and DNA polymerase I to remove the mRNA template *enzymatically* after synthesis of the first-strand cDNA. In this approach, the RNA template, now annealed to first-strand cDNA, is nicked by RNase H and degraded by the intrinsic 5'→3' exonuclease activity of *E. coli* DNA polymerase I while at the same time the 5'→3' polymerase activity of this same enzyme synthesizes the second-strand cDNA, using the 3'-OH groups from the nicked RNA as primers. Finally, as DNA polymerase is unable to seal the nicks made by RNase H, DNA ligase was included in the enzyme cocktail to repair the backbone of second-strand cDNA to maximize its stability during the subsequent cloning steps. Very cool. A major improvement to this approach was the procedure of Gubler and Hoffman (1983), in which oligo(dT)₁₂₋₁₈ rather than a (dT)-tailed vector was used to prime the first cDNA strand, followed by the synthesis of the second cDNA strand using Okayama and Berg's RNase H, DNA polymerase I, DNA ligase cocktail.



The Gubler and Hoffman method greatly simplified cDNA synthesis by eliminating the vector and concomitant circularization of the cDNA. It remains the standard format for cDNA synthesis, PCR notwithstanding. cDNA prepared by this method is usually a bit shorter than the original mRNA template due to unavoidable loss of sequence at each end of the molecule. Because cDNA synthesis in this manner does not prepare perfectly blunt ends, double-stranded cDNAs are generally incubated for a short time in the presence of T4 DNA polymerase to “polish”, i.e., create blunt ends to facilitate ligation into a vector (Fig. 17.2).

PCR-based methods

Most first-strand cDNA synthesis reactions are coupled with the PCR for the synthesis of second-strand cDNA. This is true whether the intent is to assay the abundance of specific cDNAs, for transcript mapping purposes, or for the synthesis of a traditional cDNA library. PCR is performed in a variety of formats, often featuring well-designed primers, and cleverly designed primer pairs. As the permutations are numerous and far-reaching, and the strategies diversified, to say the least, a comprehensive discussion of these approaches is best placed under the heading of “RT-PCR”, which is the subject matter of Chapter 18.

Protocol: first-strand cDNA synthesis

This protocol is just one example of how to put together a typical first-strand reaction. Typical conditions for a first-strand cDNA synthesis reaction with AMV RT are as follows: 10 mM Tris, pH 8.3; 50 mM KCl; 5 mM MgCl₂; 1 mM dNTP; 3.0 μg primer (mixture of oligo(dT) and random hexamers); 20 U RNasin; 20 U AMV reverse transcriptase; 50–1000 ng RNA; all in a final volume of 20 μl. Conveniently, the required first-strand reaction buffer accompanies

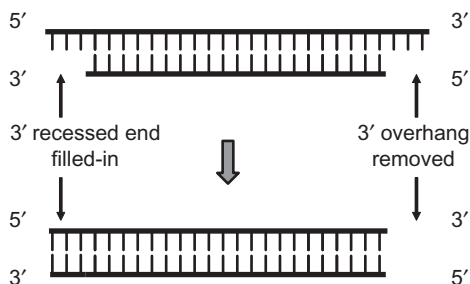


Figure 17.2 After synthesis of the second-strand, cDNA molecules often require polishing of one or both ends to maximize cloning efficiency. The enzyme T4 polymerase adds missing nucleotides in the case of a 5' overhang and degrades 3' overhangs in order to produce double-stranded molecules with two blunt ends. This is also known as “repairing” single-stranded DNA overhangs. “Repairing” or “polishing” is not necessary if the second strand was synthesized by PCR (Chapter 18).

the purchase of reverse transcriptase, minimizing or altogether eliminating the need to optimize the reaction conditions. Note that the reaction is put together in two parts, primarily to facilitate template mRNA denaturation without compromising the stability of the RNasin or the reverse transcriptase. As always, be sure to use nuclease-free techniques throughout.

1. Prepare master mix 1:

Component	Each ³
Sterile H ₂ O	5 μ l
dNTP (10 mM)	1 μ l
Random primers (2.0 μ g/ μ l)	1 μ l
Oligo(dT)12–18 (1.0 μ g/ μ l)	1 μ l
Total	8 μ l

2. Distribute 8 μ l of master mix 1 into each sample tube.
3. Add 2 μ l (50–1000 ng) of RNA⁴. Ensure an equal mass in each tube.
4. Heat tubes to 70° for 5 min.
5. Immediately place the samples on ice for at least 2 min.
6. While the RNA is being heat-denatured (step 4), prepare master mix 2:

Component	Each
Sterile H ₂ O	6 μ l
10 \times first-strand buffer	2 μ l
RNasin® (20 U/ μ l)	1 μ l
AMV RT (20 U/ μ l)	1 μ l
Total	10 μ l

7. Distribute 10 μ l of master mix 2 into each sample tube containing the denatured RNA.
8. Incubate at 25° for 5 min.

Note: This short room temperature incubation (1) facilitates primer annealing to the template, and (2) gives the reverse transcriptase an opportunity to extend the random primers and/or oligo(dT) for greater thermostability when the incubation temperature is increased in the following step. This almost always has the effect of producing much longer cDNA molecules.

³If preparing a master mix for multiple reactions, be sure to multiply the “per reaction” volumes by one more than the total number of required reactions. This is a standard procedure to ensure an adequate volume of master mix in the event of subtle micropipetting errors. This same procedure also applies to the preparation of the second-strand master mix and those used for PCR (Chapter 18).

⁴Do not add RNA directly to the master mix. Individual RNA samples are added after the master mix has been aliquoted.

9. If using a thermostable or enhanced reverse transcriptase, incubate samples per the recommendations of the manufacturer. Otherwise, incubate the sample(s) at 42° for 60 min.

Note: There are numerous RT formulations that support first-strand cDNA synthesis at elevated temperatures. These are widely available from any of a number of vendors.

10. Heat samples to 99° for 5 min to destroy the reverse transcriptase. Failure to do so may be detrimental to subsequent PCR-based procedures.

Note: The extent to which active reverse transcriptase can inhibit subsequent PCR amplification of the cDNA is a function of the ratio of enzymes and the amount of template material in the reaction tube. See Aatsinki et al. (1994) for details.

11. The newly synthesized first-strand cDNA is now ready for any one of several downstream applications. For details about RT-PCR, the reader is directed to the next chapter. For the synthesis of the second-strand cDNA for library construction, the reader is directed to the manufacturer's product guide because of great availability among the various systems on the market.

Assessing cDNA synthesis efficiency

To quantitatively determine the efficiency of reverse transcription, and as a basis of normalization among various samples, one may wish to consider spiking the first-strand reaction with an aliquot of [α -³²P] dCTP or [α -³H] dCTP; the measured amount of label incorporation is directly proportional to the mass of the newly synthesized cDNA. The mass and size distribution of radiolabeled cDNAs can be assessed by electrophoresis of an aliquot of the first- and second-strand synthesis reactions, coupled with autoradiography directly from the gel. Ideally, the exposure should show a lot of smearing toward the top of the gel: the closer the smearing is to the wells, the better. Alternatively, it is also possible, albeit a bit more cumbersome, to electrophorese an aliquot of cDNA on a gel, stain it with SYBR Green and, using image analysis software, quantify and compare the resultant smears of cDNA distributed among the samples on the gel. Whether to run an alkaline denaturing gel or not is really a matter of personal preference. One may also elect to examine the products of cDNA synthesis by briefly denaturing a small aliquot with 0.1 volume 0.1N NaOH followed by electrophoresis through a neutral 1.2% agarose gel.

In this laboratory, first-strand synthesis products are cleaned up using one of the silica-based purification products (e.g., High Pure PCR kit; Roche) after which the concentration is determined by NanoDrop spectrophotometry (Chapter 6). Finally, the size distribution of the cDNA products is examined by running a small aliquot of the sample on a 1% agarose gel and staining it with SYBR Green or SYBR Gold. A typical result is shown in Fig. 17.3.

Cloning cDNA

Upon completion of the first- and second-strand cDNA synthesis reactions, the ends of the resultant molecules are sometimes modified to yield cohesive ends

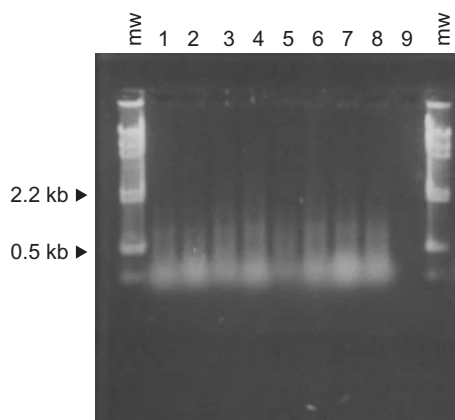


Figure 17.3 Assessment of first-strand cDNA synthesis by electrophoresis. The intensity and distribution of products within each lane is highly indicative of the quality of the cDNA and, by comparison with other samples on the same gel, the relative abundance of cDNA overall. After cleaning up the first-strand reaction, one may quantify the amount of cDNA produced by sacrificing another microliter for NanoDrop spectrophotometric analysis.

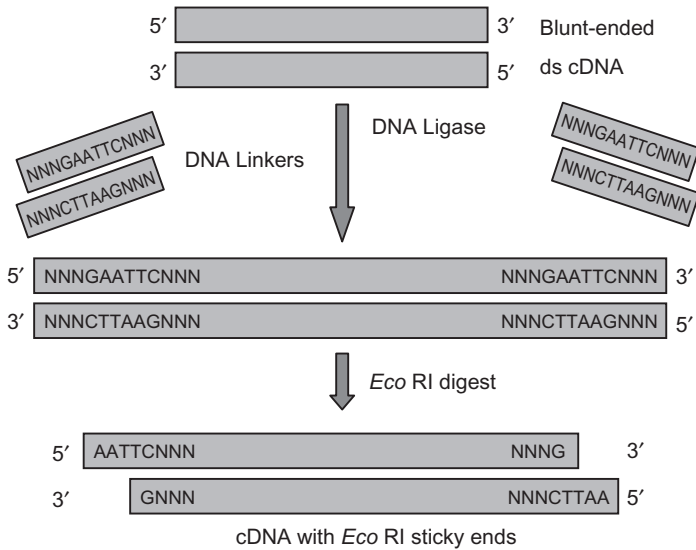
suitable for ligation to a vector (e.g., plasmid) with similar end structure. If the cDNA was primed with adapters containing rare restriction enzyme sites, then cleavage of these sequences creates “sticky ends” that will improve ligation efficiency⁵. More recently, the cloning of cDNA libraries by site-specific recombination has been described (Hartley *et al.*, 2000; Ohara and Temple, 2001; Ohara *et al.*, 2002; Invitrogen Gateway Technology).

In the absence of adapter-mediated double-stranded cDNA synthesis, blunt end ligation, while rather inefficient, can be performed to join linkers to the cDNA or ligate the cDNA directly to a vector. As shown at the top of the next page, the new restriction enzyme sites added via linker ligation are cleaved with the correct restriction enzyme to yield cohesive ends⁶. When directional cloning⁷ is required, it may be accomplished through the use of adapters with different restriction enzyme sites.

⁵In addition to linkers and adapters, a very old method for vector:insert ligation known as homopolymerically tailing still appears in the literature, albeit quite infrequently. In this procedure, double-stranded cDNA is tailed with dGTP using terminal deoxynucleotidyl transferase (terminal transferase, for short). The vector is also tailed, but with dCTP. The resulting poly(dC) and poly(dG) tracts, obviously complementary, are annealed and then joined together with DNA ligase.

⁶If the restriction enzyme used to generate cohesive ends also cuts an internal recognition site, the cDNA will be diminished in size. To preclude this possibility, one strategy is to methylate the cDNA (e.g., with *Eco* RI methylase) prior to linker ligation and restriction enzyme digestion. In this example, methylated *Eco* RI sites in the body of the cDNA would be resistant to *Eco* RI cutting.

⁷A different restriction enzyme site on each end of double-stranded cDNA permits ligation of an insert in a desired orientation. This strategy is often used when the cloned cDNA is to be expressed.



Virtually all cDNA synthesis and cloning systems provide the user with linkers and/or adapters to facilitate ligation into a vector of choice. When using PCR, the primers must be designed in advance to provide an end structure compatible with the intended method of cloning. Finally, many of the intrinsic cloning difficulties described here can be avoided altogether by using the TOPO Cloning approaches described in Chapter 18.

Ligation considerations

Maximum ligation efficiency is favored when the molecules to be ligated, in this case vector and insert, are present in a 1:1 molar ratio. In the synthesis of a library, a near 1:1 ratio of vector to insert will favor completeness in cloning. In the event that the generation of recombinant clones is problematic, ligation optimization is accomplished by calculating, to the extent possible, the number of picomole (pmol) ends involved in the reaction, and then increasing or decreasing the mass of one of the “participants” in other to achieve 1:1 vector:insert ends. One must calculate pmol ends for both the insert and the vector, and then equalize them in the ligation reaction.

$$\text{pmol ends} = \frac{\mu\text{g DNA}}{660 \times \text{number of bases}} \times 2 \times 10^6$$

where

pmol ends = number of molecule ends available to participate in the ligation reaction
 μg DNA = mass of the DNA (either the insert or the vector)

660 = average molecular weight of a nucleotide base pair

Number of bases = the length of the DNA (either insert or vector)

2×10^6 = conversion factor from μg (10^{-6}) to pmol (10^{-12}) and the fact that each molecule has two ends capable of ligation.

Thus, 575 ng of a DNA sequence that is 852 bp long contains 2.045 pmol ends:

$$\text{pmol ends} = \frac{0.575 \mu\text{g}}{660 \times 852} \times 2 \times 10^6$$

Assuming, for cloning purposes, that this DNA is intended to be an insert, the investigator would also need to determine the mass of the intended vector needed to give an equivalent number of pmol ends, thereby ensuring an equimolar concentration of each in the ligation reaction.

This calculation is very simple and straightforward when an insert of uniform length is intended to be cloned. The situation is quite a bit more complicated when a mixed population of cDNAs is to be cloned because the cDNAs are all different lengths, and some are much more abundant than others. In this case, the most worthwhile strategy is to vary the insert:vector ratio in a series of ligations, an example of which might be:

1:0.25, 1:0.5, 1:0.75, 1:1, 1:1.25, 1:1.5, 1:1.75, 1:2

At the conclusion of the ligation reaction, the now recombinant molecules are used to transform competent *E. coli* cells⁸, most often using a chemical transformation process and ideally with competent cells that have been purchased (e.g., Promega).

If, however, one is simply attempting to clone a single type of molecule, rather than a library, ligation ratios are much less of a concern. After bacterial transformation and plating onto selective media (e.g., 50 $\mu\text{g}/\text{ml}$ ampicillin or 12.5 $\mu\text{g}/\text{ml}$ tetracycline), if even a single colony appears and is found to possess the desired insert, then the experiment is a success. Since false positives do occur, it is best to have at least a few candidate clones to examine for the presence of the recombinant vector. The take home lesson: a perfect 1:1 ratio is often not a critical parameter for simple cloning applications.

⁸cDNA library construction is most efficient when cloning into one of the popular λ bacteriophage vectors rather than into plasmids. Because λ bacteriophage plaques are very small, the library can be plated at a much higher density than a corresponding number of bacterial transformants. False positives are also much less frequent when screening λ libraries, thereby conferring enhanced sensitivity and resolution, compared to colony hybridization. After screening a λ library, the insert can be rescued from putative positives by PCR for subsequent characterization.

Enzymes used for ligation

The master enzymes for ligation are DNA ligase and topoisomerase. For most routine cloning experiments T4 DNA ligase is in widespread use. The natural source of this enzyme is T4 (virus)-infected *E. coli*. Like all DNA ligases, this enzyme creates phosphodiester linkages between 5' phosphate groups on one molecule and 3' hydroxyl groups on a proximal molecule. The action of DNA ligase might best be thought of as “molecular paste” which can either undo the action of “molecular scissors” (restriction enzymes) or that can create novel DNA sequences by cleverly planning a series of ligation reactions. DNA ligase has the ability to stably join DNA molecules with cohesive (sticky) ends or blunt ends. Another useful activity associated with ligases is the ability to repair nicks in one of the two strands of dsDNA, dsRNA, and DNA:RNA hybrid molecules. Ligation reactions often require incubations ranging from 2–24 hours. T4 RNA ligase is likewise available for the ligation of both ssDNA and ssRNA.

More recently, TOPO[®] Cloning (Invitrogen) has become a widely respected alternative approach for cloning dsDNA. In this system, the insert of interest is combined with a linearized TOPO vector to which the enzyme topoisomerase I has already been joined. These “activated” vectors are therefore ready for immediate use and do not exhibit self-ligation. This method of cloning is rapid and efficient, often rendering >95% cloning efficiency in only a few minutes at room temperature. There are several TOPO vectors that support a variety of cloning formats. Upon ligation, the topoisomerase I enzyme is released from the now recombinant plasmid construction. Bacterial transformation proceeds as usual. This approach to cloning is described in greater detail in Chapter 18.

Applications

The ability to synthesis cDNA which is representative of the starting RNA template is an important aspect of molecular biology. The emphasis in cDNA synthesis and cloning is on active gene expression; by its very nature, the synthesis of cDNA creates a permanent snapshot of gene expression at the moment the RNA was harvested from the biological source. Unlike intrinsically labile RNA, cDNA can be archived for years in the freezer and when constructed with appropriately characterized ends, the cDNA can be propagated indefinitely. cDNA has very important clinical implications, too. Retroviral detection, after all, involves the conversion of a viral RNA chromosome into cDNA, which may then be amplified several million-fold by PCR or other amplification procedures. Finally, appropriately designed cDNA synthesis experiments are quite useful for mapping the transcription start site (TSS) of an RNA molecule, mapping the 3' end of a transcript, and for the discovery of alternative splicing patterns associated with the RNA molecules produced from a single genetic locus. These and other applications are presented in greater detail in the following chapter.

References

- Aatsinki, J.T., J.T. Lakkakorpi, E.M. Pietila, and H.J. Rajaniemi. (1994). A coupled one-step reverse transcription PCR procedure for generation of full-length open reading frames. *BioTechniques* **16**, 282–285.
- Bassett, C.L., M.L. Nickerson, R.A. Cohen and M.S. Rajeevan. (2000). Alternative transcription initiation and novel post-transcriptional processing of a leucine-rich repeat receptor-like kinase gene that responds to short-day photoperiodic floral induction in morning glory (*Ipomoea nil*). *Plant Mol. Biol.* **43**, 43–58.
- Bassett, C.L, M.L. Nickerson, R.E. Farrell, Jr., and M. Harrison. (2004). Multiple transcripts of a leucine-rich repeat receptor kinase from morning glory originate from different TATA boxes in a tissue-specific manner. *Mol. Gen. Genomics* **271**(6), 752–760.
- Chenchik, A., Y.Y. Zhu, L. Diatchenko, R. Li, J. Hill, and P.D. Siebert. (1998). Generation and use of high quality cDNA from small amounts of total RNA by SMART PCR. In *Gene Cloning and Analysis of RT-PCR* (P.D. Siebert and J. Larrick, Eds). BioTechniques Books, Natick, MA. pp. 305–319.
- Efstratiadis, A., F.C. Kafatos, A.M. Maxam, and T. Maniatis. (1976). Enzymatic *in vitro* synthesis of globin genes. *Cell* **7**, 279–288.
- Gubler, U. and B.J. Hoffman. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.
- Hartley, J.L., G.F. Temple, and M.A. Brasch. (2000). DNA cloning using *in vitro* site-specific recombination. *Genome Res.* **10**, 1788–1795.
- Matz, M., D. Shagin, E. Bogdanova, O. Britanova, S. Lukyanov, L. Diatchenko and A. Chenchik. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res.* **27**, 1558–1560.
- Murakawa, G.J., J.A. Zaia, P.A. Spallone, D.A. Stephens, B.E. Kaplan, R.B. Wallace and J.J. Rossi. (1988). Direct detection of HIV-1 RNA from AIDS and ARC patient samples. *DNA* **7**, 287–295.
- Ohara, O., T. Nagase, G. Mitsui, H. Kohga, R. Kikuno, S. Hiraoka, Y. Takahashi, S. Kitajima, Y. Saga, and H. Koseki. (2002). Characterization of size-fractionated cDNA libraries generated by the *in vitro* recombination-assisted method. *DNA Res.* **9**(2), 47–57.
- Ohara, O. and G. Temple. (2001). Directional cDNA library construction assisted by the *in vitro* recombination reaction. *Nucleic Acids Res.* **29**, e22.
- Okayama, H. and P. Berg. (1982). High efficiency cloning of full-length cDNA. *Mol. Cell. Biol.* **2**, 161–170.
- Rougeon, F. and B. Mach. (1976). Stepwise biosynthesis *in vitro* of globin genes from globin mRNA by NA polymerase of avian myeloblastosis virus. *Proc. Natl. Acad. Sci. USA* **73**, 3418–3422.
- Sellner, L.N. and G.R. Turbett. (1998). Comparison of three RT-PCR methods. *BioTechniques* **25**, 230–234.

18 RT-PCR: A Science and an Art Form

Rationale	385	Non-PCR methods for confirming PCR-derived data	423
PCR – an overview	386	Related techniques	424
RT-PCR – general approach	391	5' RACE PCR	424
PCR carryover prevention	395	5' RLM-Race	426
Laboratory design	395	3' RACE PCR	426
Procedural methods	396	Nested PCR	428
Aerosol-resistant tips	397	Long-range PCR	429
Uracil-N-glycosylase	398	Single-cell PCR	431
Primer design	398	Splinkerette PCR	433
Basic rules	402	The hunt for alternative transcription start sites	434
T_m considerations	406	Protocol: first-strand cDNA synthesis	435
Estimating T_m	406	Protocol: PCR amplification of cDNA	436
Precision T_m calculations	407	Cloning PCR products	437
On-line resources	407	Protocol: a-tailing of blunt-end PCR products	438
Multiplex primer design	407	Protocol: TA cloning ligation reaction	438
Optimization procedures	408	TOPO® Cloning	440
Thermostable polymerases	411	Other amplification procedures	441
Positive controls	413	linear RNA amplification (Eberwine process)	441
Negative controls	413	Strand displacement amplification	442
Hot start PCR	413	Nucleic acid sequence based amplification (NASBA)	443
Locked nucleic acids	415	Ligase chain reaction	443
Touchdown PCR	416	References	443
Internal controls	416		
The word on transcription controls	419		
Analysis of PCR products	420		
RT-PCR quality control points	421		

Rationale

Without a doubt, the revolutionary polymerase chain reaction (PCR) has influenced profoundly all of molecular biology, even to the most fundamental level of *how* a particular experimental problem ought to be approached. RT-PCR is that technology by which RNA molecules are converted into their complementary

DNA (cDNA) sequences by any one of several reverse transcriptases, followed by the amplification of the newly synthesized cDNA by standard PCR procedures. Because of the role of reverse transcriptase (RT) in the synthesis of first-strand cDNA, this approach to studying gene expression is universally known as RT-PCR. Interestingly, this technique was originally referred to as RNA PCR, which was something of a misnomer. Because of the speed and efficiency of PCR, RT-PCR is now the preferred method for the synthesis and analysis of cDNA.

The utility of RT-PCR is readily apparent when compared to the traditional methods of RNA analysis, which pale by comparison. The emphasis of RT-PCR, as well as all methods involving cDNA, is on transcriptionally active loci: a transcript must be present in a cell lysate to support reverse transcription and amplification by PCR. Genes that are transcriptionally silent, therefore, are not assayable if no template exists in the reaction tube. Given the extreme sensitivity of PCR, it is now commonplace to detect and quantify transcripts present in extremely low abundance.

PCR – an overview

Unquestionably, the widespread implementation of the polymerase chain reaction has fundamentally influenced all aspects of basic research in the life sciences. The popularity of many classical labor-intensive methods of vintage molecular biology (circa 1985) continue to fade in favor of the faster, more discriminating method of what is precisely defined as a “primer-mediated, enzymatic amplification of specific genomic or cDNA sequences” (Saiki *et al.*, 1985; Mullis *et al.*, 1986; Mullis and Faloona, 1987). That minute quantities of template material can be amplified several millionfold in a matter of a few hours has unleashed enormous potential in the study of infectious disease, gene mutations, forensics, ontological relationships among members of gene families, and transcriptional activity in cells and tissues, using RNA as a parameter of gene expression. One should not, however, consider abandoning the classical methods entirely, because of several intrinsic shortcomings of the polymerase chain reaction, namely:

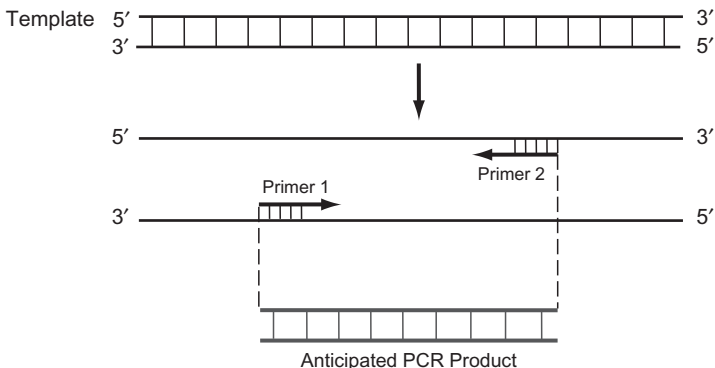
1. In order for the reaction to function properly (at all), the investigator must have direct or indirect knowledge of appropriate primer sequences.
2. A primer, when base-paired to the template, must not have a mismatch at its 3'-OH terminus.
3. The template for the reaction (i.e., the target) must be clean enough to facilitate primer annealing and subsequent elongation.
4. The various components of the reaction cocktail must be optimized to yield unambiguous data.
5. The length of a PCR product is generally much shorter than the length of the original template. In the case of RT-PCR, the resulting product can be purified and used as a Northern analysis probe to identify the size of the mRNA template.
6. PCR is a biased amplification procedure. Some sequences amplify with greater efficiency than other sequences.
7. From a relative abundance perspective, PCR tends to distort the true abundance relationships among cDNAs.

It is the opinion of this Author that the polymerase chain reaction should not be viewed as a cure-all for the numerous difficulties that frequently frustrate both seasoned and novice molecular biologists. Further, one should not attempt implementation of PCR to the exclusion of other molecular methods, even if one is reasonably proficient with basic isolation, handling, and storage methods for nucleic acids. More than one student of molecular biology has mistakenly believed that the immediate implementation of PCR will solve all problems in the lab. False.

The mechanics of the polymerase chain reaction have been presented and discussed in resources too numerous to list completely; for excellent reviews the reader is directed to Dieffenbach and Dveksler (2003); Innis, *et al.* (1999); and McPherson and Hames (1995). For the convenience of the reader, a brief overview of PCR is presented here.

PCR is dependent, first and foremost, on the primers included in the reaction. Primers, in this context, are short, single-stranded DNA sequences, also known as oligonucleotides, which has been artificially synthesized to the specifications of the investigator. The essential qualities of a primer are:

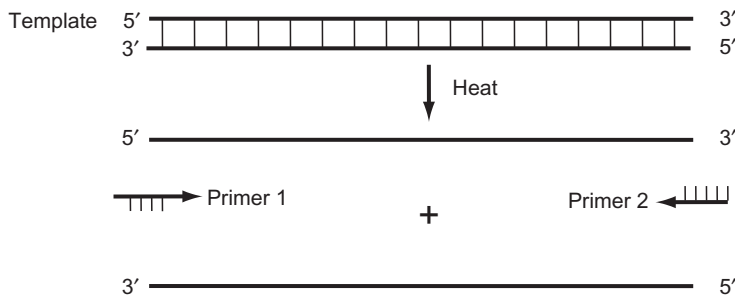
1. Primers are short, usually less than 30 bases. A primer that is 20–25 bases long is usually ideal for routine applications.
2. Primers are single-stranded.
3. Primers manifest a free 3'-OH group that is available to support the addition of nucleotides by a DNA polymerase. The 5' end of a primer, in contrast, may carry all types of exotic modifications, without compromising its specificity.
4. The 3' terminal base, and often the penultimate base, must be perfectly base-paired to the template in order to support primer extension. In contrast, the 5' end of a primer needs to be base-paired to the template at all.
5. The primers must show at least some degree of complementarity with the template, i.e., the experimental genomic DNA or cDNA that the investigator wishes to amplify.
6. Primers are used in pairs. The primers must base-pair to opposite strands of the heat denatured template in such a way that their respective 3' ends face each other. Thus, the anticipated PCR product is that which is framed by the 5' termini of a primer pair.



It is imperative that the primers be used at an equimolar concentration, so that each strand of template is amplified symmetrically, hence the terminology “symmetric PCR” (Saiki *et al.*, 1985). This is the norm. If the primer concentrations are different, an asymmetric reaction (Gyllensten and Erlich, 1988) may result in which one of the target strands is amplified preferentially. Depending on the way the primers are designed, the polymerase chain reaction will support, in addition to symmetric PCR and asymmetric PCR reactions, a variety of other amplification formats, including inverse PCR (Ochman *et al.*, 1988; Triglia *et al.*, 1988), RT-PCR (Veres *et al.*, 1987), and rapid amplification of cDNA ends, also known as 5'- and 3' RACE (Frohman *et al.*, 1988; Frohman, 1995), to name but a few variations on this theme.

The polymerase chain reaction is divided into cycles, with each cycle typically consisting of three components in which the sample experiences substantial temperature changes. The orderly application of these functionally distinct components, typically in the form of 25–30 cycles, results in the exponential amplification of template molecules (Table 18.1). The three distinct components of each cycle are as follows:

1. *Denaturation*, the goal of which is high temperature dissociation of double-stranded template molecules into their constituent single strands. This includes denaturation of any secondary structures that may have formed within or among the primers¹. This is necessary to facilitate annealing of primers to the template in the next step. Template denaturation is an absolute requirement if any PCR product is to be observed.



The precise temperature and time required for denaturation is a direct function of the complexity of the target as well as the G+C content. Thus, the more knowledge one has as to the nature of the template, the easier any necessary optimization will be later on. In general, genomic DNA samples require prolonged denaturation to ensure complete dissociation of the template into its constituent single strands. Failure to produce PCR product from a genomic template is very often the result of incomplete target denaturation from the start, which greatly impedes the efficiency of subsequent cycles. Typically, initial denaturation is conducted at 94° for as little as 10–30 s for short, linear DNA sequences, and as long as 10 minutes for higher complexity genomic DNA.

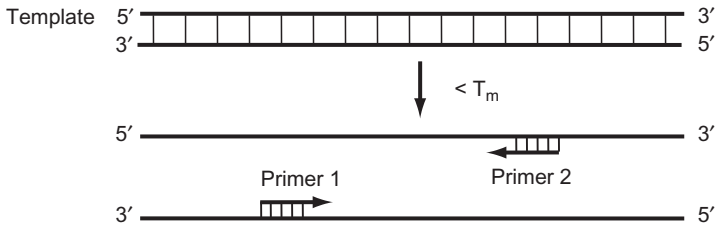
¹See the section on primer design for suggestions on how to avoid intramolecular and intermolecular base pairing of primers.

Table 18.1 PCR Theoretical Exponential Amplification

Cycle number	Relative product mass	Base 2 exponent (2^n)
1	2	2^1
2	4	2^2
3	8	2^3
4	16	2^4
5	32	2^5
6	64	2^6
7	128	2^7
8	256	2^8
9	512	2^9
10	1,024	2^{10}
11	2,048	2^{11}
12	4,096	2^{12}
13	8,192	2^{13}
14	16,384	2^{14}
15	32,768	2^{15}
16	65,536	2^{16}
17	131,072	2^{17}
18	262,144	2^{18}
19	524,288	2^{19}
20	1,048,576	2^{20}
21	2,097,152	2^{21}
22	4,194,304	2^{22}
23	8,388,608	2^{23}
24	16,777,216	2^{24}
25	33,554,432	2^{25}
26	67,108,864	2^{26}
27	134,217,728	2^{27}
28	268,435,456	2^{28}
29	536,870,912	2^{29}
30	1,073,741,824	2^{30}
31	2,147,483,648	2^{31}
32	4,294,967,296	2^{32}
33	8,589,934,592	2^{33}
34	17,179,869,184	2^{34}
35	34,359,738,368	2^{35}

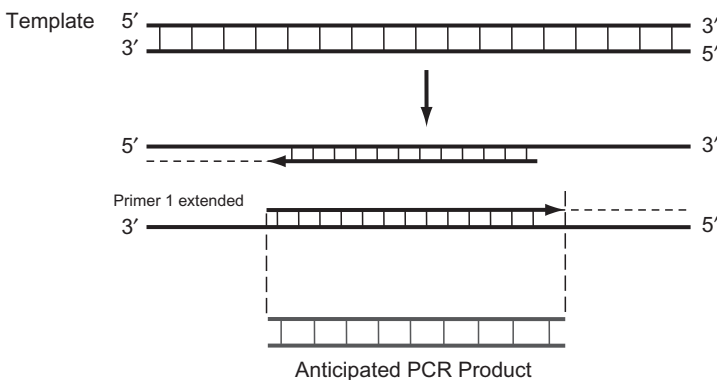
The theoretical accumulation of PCR product for any number of cycles is 2^n , where n = number of cycles completed. In reality, however, the actual amplification of template is much less, due to the depletion of available primer and dNTPs, as well as the cumulative negative impact of repeated sample denaturation (92–95°) on the polymerase, and other issues. Product accumulation is exponential in the earlier cycles and subsequently demonstrates what is known as a “plateau effect” in later cycles. This means that although product accumulation may continue, it does so in a non-exponential manner and only as long as the polymerase remains active.

2. *Primer annealing*, the goal of which is the hybridization of the primers to complementary sites on the template. In so doing, the primers provide the 3'-OH group needed to support DNA synthesis in the next step. When a set of primers anneals to the template the 5' ends of the primer frame, or define the size of, the PCR product in Cycle 1 (Cycle 2 for RT-PCR). For annealing to occur, the temperature inside the reaction tube should not exceed the melting temperature (T_m) of either of the oligonucleotide primers. For this reason, the annealing temperature (T_a) is usually at least 1° to 2° below the lowest calculated T_m of the primers involved, although exceptions abound.



Moreover, primer annealing will not take place unless the region of the target material containing sequences complementary to the primers has been sufficiently denatured². For these reasons, the annealing temperature is the most variable cycling component in PCR, and is a direct function of the primer sequences. Typical annealing temperatures range from 50° to 68° for 10–60 s.

3. *Primer extension*, the goal of which is synthesis of the desired PCR products by primer extension along the template molecule to which the primers are base-paired. This is the heart of PCR product synthesis. The inclusion of one or more thermostable DNA polymerases, e.g., *Taq* polymerase, from the thermophilic microbe *Thermus aquaticus*, supports primer extension at elevated temperatures. These enzymes can withstand the high temperatures associated with the denaturation component, as well as the primer extension component, most often 68° to 72°.



In cycles one and two, primer extension occurs beyond the region defined by the 5' end of the primer on the opposite strand, resulting in the production the

²See the section on optimization procedures for suggestions.

“long” and “intermediate” products of PCR. By cycle three the reaction is producing the “short” product, that is, the desired product that accumulates exponentially in the reaction tube. Intermediate product is produced throughout the reaction, but only accumulates arithmetically, and is really a non-issue with respect to the quantitiveness of the assay (Table 18.2). Non-specific product, however, is another issue entirely. Further, thermostable polymerases make fewer mistakes at elevated temperatures, though it is important to note that not all thermostable enzymes have a proofreading activity. Because the error rate of the *Taq* polymerase can be significant, many of the newer protocols call for the use of enzyme blends, that is, two different thermostable enzymes functioning in a permissive manner, to increase the yield and fidelity of the amplification process.

In summary, the polymerase chain reaction consists of the repeated application of the three steps described previously: template denaturation, primer annealing, and primer extension. While the PCR product yield is quite unimpressive for the first several cycles, the later cycles (cycle 20 and beyond) are responsible for generating the bulk of the usable product. Regardless of the name given to a PCR-based technique, the fundamental replication of these steps will always apply.

RT-PCR – general approach

It is easy to see how PCR amplification is accomplished starting with genomic DNA or even vector DNA as a template. The same principles apply with respect to the amplification of RNA, with an additional step required to first convert RNA “templates” into cDNA. Given the complex nature of any assay designed to provide a clearer understanding of gene regulation, conditions governing the success or failure of RT-PCR require careful forethought coincident with thorough empirical determination.

The historical focus associated with cDNA has centered on the synthesis of full-length or *representative* cDNA molecules, and the extent to which this can be accomplished is directly related to the quality of the RNA starting material, the method of priming for the first-strand cDNA, and the ability of reverse transcriptase to plow through secondary structures or RNA hairpins. Because of the unsurpassed sensitivity and amplification potential of PCR, the quantity of starting RNA is less important in terms of its ability to be PCR-amplified, by comparison with more traditional cloning methods. Moreover, because PCR primers “frame” the domain that will be amplified, this method is far more tolerant of partially degraded RNA than are traditional cDNA synthesis techniques and Northern blot analysis.

RT-PCR is a two-step process. First, purified RNA is reverse transcribed by RT (e.g., MMLV, AMV, Superscript®) via an appropriate method for priming³. Second, first-strand cDNA is amplified using some variant of PCR. As with

³In this reaction, which is first-strand cDNA synthesis, one may utilize traditional downstream primers, such as oligo(dT)_{12–18}, random primers, or gene-specific primers. The net result is the synthesis of stable, single-stranded cDNA molecules capable of acting as a template for amplification by PCR.

Table 18.2 PCR Product Inventory

Relative product distribution by type				
Cycle number (n)	Long product (constant)	Intermediate product (2 ⁿ -2)	Short Product (2 ⁿ -2 ⁿ)	Total (2 ⁿ)
1	2	0	0	2
2	2	2	0	4
3	2	4	2	8
4	2	6	8	16
5	2	8	22	32
6	2	10	52	64
7	2	12	114	128
8	2	14	240	256
9	2	16	494	512
10	2	18	1,004	1,024
11	2	20	2,026	2,048
12	2	22	4,072	4,096
13	2	24	8,166	8,192
14	2	26	16,356	16,384
15	2	28	32,738	32,768
16	2	30	65,504	65,536
17	2	32	131,038	131,072
18	2	34	262,108	262,144
19	2	36	524,250	524,288
20	2	38	1,048,536	1,048,576
21	2	40	2,097,110	2,097,152
22	2	42	4,194,260	4,194,304
23	2	44	8,388,562	8,388,608
24	2	46	16,777,168	16,777,216
25	2	48	33,554,382	33,554,432
26	2	50	67,108,812	67,108,864
27	2	52	134,217,674	134,217,728
28	2	54	268,435,400	268,435,456
29	2	56	536,870,854	536,870,912
30	2	58	1,073,741,764	1,073,741,824

The accumulation of PCR products in a typical symmetric reaction is distributed among three forms, namely, the long product, the intermediate product, and the short product. Examination of the “short product” and “total” columns reveals the rapid, exponential accumulation of product in a reaction tube, the vast majority of which is the short product. The intermediate product continues to accumulate, but only does so arithmetically. The actual amount of PCR product present at the end of the reaction is also a direct function of the amplification efficiency, which is always less than the 100% efficient model depicted here.

all multistep procedures, the outcome of the experiment is determined by the efficiency of each individual reaction. Along these lines, an important issue is whether the process of RT-PCR is performed in one’s laboratory in one tube or in two. The original method, which remains in use today, has been to prepare

cDNA in one reaction and then remove aliquots for subsequent amplification in a second tube using a completely different reaction chemistry. A more recent innovation, known as single-tube or one-tube RT-PCR, has streamlined the assay of specific transcripts and is particularly suited for clinical and diagnostic applications. In the one-tube reaction format, reverse transcription of the template RNA and the amplification of the resulting cDNA occur in a single tube.

There are several variants on the one-tube RT-PCR theme. In the original one-tube format, the thermostable enzyme *rTth* polymerase is used to create first-strand cDNA in a Mn^{2+} -dependent reaction in a bicine⁴ buffering environment. Then, chelation of the Mn^{2+} and addition of a Mg^{2+} buffer coaxes the same enzyme to support PCR amplification. While rather innovative at the time of its development, the assay did require opening the tube in order to add reagents after making cDNA but before amplifying it. In the second case, all necessary cDNA synthesis *and* PCR components could be premixed from the onset (Aatsinki *et al.*, 1994). One method involves mixing *Taq* polymerase and AMV reverse transcriptase together, while another variant involves maintaining *Taq* polymerase in an inactivated state by association with an anti-*Taq* antibody. Reverse transcription proceeds as usual. Then, heating to 95° destroys the reverse transcriptase and the antibody, thereby liberating the *Taq*, and PCR then proceeds as usual. Obviously, much was gained when this one-tube reaction format evolved. The key advantage of this latter method is that once the reaction is assembled and the tube closed, no additional components are required nor is there any reason to open the tube until it is time to analyze the PCR products by electrophoresis. In the case of real-time PCR, quantification will not require a gel at all. In yet another format, a KCl and $(NH_4)_2SO_4$ buffer system ensures high specificity over a dynamic range of temperatures and Mg^{++} concentrations, minimizing the need for optimization. Improvements such as these, better instrumentation, the fact that many researchers now understand the reactions better, and other innovations have drastically reduced the required mass of input RNA, sometimes to sub-nanogram quantities. At levels such as these, single cell PCR becomes a very real possibility for the investigator.

The first-strand cDNA synthesis reaction produces an RNA:DNA hybrid. In the second reaction, the mechanics of the polymerase chain reaction are responsible for the displacement of the original RNA template: at the onset of cycle 1 the heat denaturation step and quasi-alkaline environment produced by the PCR reaction buffer cause RNA:DNA strand denaturation and RNA alkaline hydrolysis, respectively⁵. This clears the way for sequence-specific upstream primer annealing needed to support synthesis of the second-strand.

⁴KCl and $(NH_4)_2SO_4$.

⁵The primary advantages of RT-PCR over traditional methods of cDNA synthesis are speed and sensitivity. Unlike the still popular RNase H-mediated method of second-strand cDNA synthesis (Gubler and Hoffman, 1983), alkaline hydrolysis of the original RNA template and immediate synthesis of the second-strand cDNA eliminates various clean-up steps and other tedious manipulations. By designing fairly non-specific primers, for example oligo (dT) and combinations of random hexamers or nonomers, entire libraries can be synthesized using PCR which can then be archived for years.

The primer extension component of the first cycle will, in turn, generate the second-strand cDNA. After cycle 1 is completed, the amplification of both strands of the newly synthesized cDNA will proceed symmetrically (Fig. 18.1).

In the days before PCR, second-strand cDNA synthesis was initiated by the addition of the enzyme DNA polymerase, often without prior removal or inactivation of the reverse transcriptase used to make first-strand cDNA. DNA polymerase and reverse transcriptase working together have an enhanced ability to plow through secondary structures that can compromise the length of the resulting double-stranded cDNA. With respect to PCR, however, with the exception of the Mn-dependent *rTth* and *Taq* activities, failure to heat-inactivate reverse transcriptase may interfere with subsequent applications (Kawasaki, 1990); this effect may be partially overcome by increasing the ratio of *Taq* polymerase to reverse transcriptase (Sellner *et al.*, 1992, Sellner and Turbett, 1998).

One might think that the assay of transcripts by RT-PCR should be just as easy as amplification of genomic DNA because cells make multiple copies of the mRNA, even of the low abundance variety. Just the opposite is true, however, because the investigator is at the mercy of the reverse transcriptase. The potential

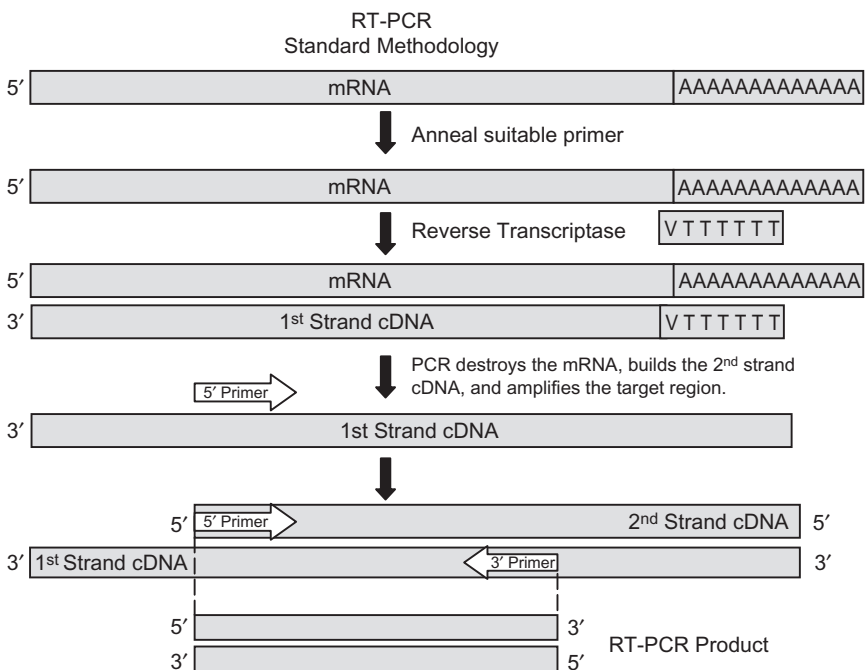


Figure 18.1 RT-PCR. mRNA is converted into first-strand cDNA by the action of reverse transcriptase. The mRNA template is then destroyed by the mechanics of the PCR component of this technique. The second-strand cDNA is synthesized in PCR cycle 1, followed by enzymatic amplification of double-stranded cDNA beginning with cycle 2. Numerous alternative priming strategies are possible. See text for details.

exists for significant variability in reverse transcriptase efficiency among multiple samples, even when pipetting from the same master mix. Consequently, normalization of samples becomes all the more important in RT-PCR, and the efficiency of first-strand cDNA synthesis is one of the most important determinants in the success or failure of this method. For this reason, it is strategically better to make a large cDNA pool from which aliquots may be drawn for individual applications rather than repeating the same cDNA synthesis reaction over and over.

PCR carryover prevention

The major advantage of PCR is also the major disadvantage. The amplifying power associated with this process occasionally causes problems when (1) *non*-specific products are coamplified, often in competition with the intended template; and (2) when products from a previous reaction manage to contaminate subsequent reactions. In the case of the latter, amplicon contamination can occur at any of a number of different levels. It is imperative that steps be taken to prevent amplicon contamination, and that unambiguous negative controls be performed along with each set of experimental samples. False-positive data are widely observed as a consequence of uncontrolled amplicon release coupled with the sensitivity of PCR.

Invariably, failure to address the potential for amplicon contamination will, sooner or later, lead to full-blown contamination in the lab. Suddenly, negative controls will show up as positives, the reliability index of quantitative assays will approach zero, clinical samples from healthy individuals will show up as false positives, and any manufacturing process that involves a PCR product or byproduct will come to a grinding halt. These are just a few of the dire, far-reaching consequences associated with contamination.

Amplicon contamination is also known as carry-over contamination because the chief source of the contamination is a PCR product which “carried over” from a previous experiment. There are specific procedures, equipment, and laboratory policies that, if followed, should prevent carry-over contamination from occurring in the lab. These topics can be broadly classified into the four categories: laboratory design and workflow, various procedural methods, the use of aerosol-resistant tips and positive-displacement micropipettors, and the appropriate use of the enzyme uracil-N-glycosylase. The procedures and policies described below apply to low throughput and high throughput laboratories alike.

Laboratory design

The major objective in laboratory design is to segregate PCR products from all other areas involved in pre-amplification activities, such as template preparation, preparation of stock solutions and master mixes, and PCR reaction assembly. A typical lab layout is shown in Fig. 18.2. While it is clearly desirable to have adequate laboratory square footage available to maintain separate

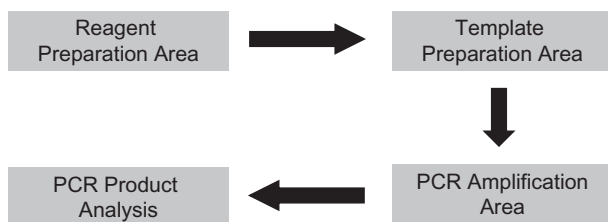


Figure 18.2 Preferred arrangement of a work area or laboratory utilizing PCR. Reagents, templates, and other materials move only in the direction indicated by the arrows, and are *never* returned to a previous station. Moreover, a set of micropipettors and laboratory coats should be designated for each station and should not travel between one station and the next. Unidirectional movement of stock solutions and sample tubes is a critical for preventing amplicon contamination.

rooms for various pre- and post-PCR tasks, this is not always possible in smaller laboratories. Instead, to minimize the inherent dangers of amplicon contamination (PCR carryover), it is best to at least define areas within the lab where template alone will be handled, and then identify areas exclusively for reagent preparation and PCR cycling. Further, investigators should be very cognizant of which tubes are moved into each area, and in which direction.

In general, four areas should be designated. The movement or flow of materials should always be from the area with the cleanest or most highly purified materials (reagent preparation), to a lesser purity area (template preparation), to the final area containing the least pure or dirtiest materials (the location of the thermal cycler) and then finally an area where electrophoresis is performed (product analysis), if applicable. Each area should have a dedicated set of micropipettors, lab coats and other instrumentation, including microfuges. Nothing should be moved from one area to the next, except for the sample tubes and aliquots taken from stock solutions that are to be used for the experiment at hand. It is wise to at least organize the lab for unidirectional movement, which will reduce the incidence of carryover contamination. When laboratory space is at a premium, it is most judicious to run gels or otherwise perform PCR product analysis as far away as possible from the other three areas. Keep in mind also that the additional steps involved in RT-PCR (format-dependent) afford greater opportunities for amplicon contamination of the stock solutions and downstream reactions, compared to the normal peril associated with non-RT-PCR. In order to address this issue, the investigator is well advised to perform the reverse transcriptase reaction in the sample preparation area, as opposed to the PCR amplification area, if such a designation has been made in the lab.

Procedural methods

It is incumbent upon the head of the laboratory to establish, update, and enforce standard operating procedures associated with the use of PCR. While some of these procedures may be taken for granted by the more senior members

of the lab, appropriate procedures and behaviors may not be immediately intuitive to new personnel. Each new lab member should be required to not only go through an orientation of sorts to familiarize that person with the proper and safe operation of equipment, but should also be given a series of simple experiments to perform in order to assess their competency performing PCR. Beyond learning solid laboratory skills, there are other approaches that may be helpful in minimizing carry-over contamination. In some laboratories, the thermal cycler is maintained in a laminar flow hood (the type used for tissue culture). At some predetermined interval, perhaps at the end of each day or each week, the block into which the sample tubes are placed is briefly UV-irradiated in order to cross-link any amplicons, thereby eliminating their ability to cause experimental turmoil. Alternatively, a hand-held UV device can be mounted over the block for the same purpose. As always, be sure to avoid any eye or skin exposure to UV light.

With respect to master mixes, it is acceptable to UV-irradiate the water and 10× reaction buffer prior to the addition of the template, primers, or enzyme, in order to control amplicon carryover. While this measure is not foolproof because enzyme, primers, and template are added postirradiation, it may help. One might also consider the use of isopsoralen which, when exposed to long-wave UV light at the end of the PCR, will form covalent adducts involving pyrimidines in the DNA products, thereby preventing their future functionality as a template molecule (Isaacs *et al.*, 1991; Fahle *et al.*, 1999). Isopsoralen is added to the reaction tube(s) prior to the cycling and has no apparent effect on amplification efficiency. UV irradiation occurs after cycling is complete but before the tube is opened.

Aerosol-resistant tips

Aerosol contamination can be a problem of immense magnitude if it occurs in the lab and is probably the most common means by which carryover contamination occurs. Aerosols are formed when a liquid is drawn up into a micropipette tip. The vacuum that is created by the micropipettor can draw amplicons into the bottom of the micropipettor shaft, thereby making the instrument itself a vector for contamination. Aerosol-resistant tips, also known in some circles as barrier tips (Fig. 18.3a), contain a hydrophobic polyethylene filter positioned within the tip that not only effectively prevents aerosols from moving up past the filter and into the shaft, but also prevents anything that might already be in the shaft from coming down and contaminating the solution in the tip. Barrier tips are supplied sterile and are not to be autoclaved. Alternatively, the use of positive displacement micropipette tips (Fig. 18.3b) may also be quite helpful in this regard, though positive displacement tips are more expensive than barrier tips. Regardless of the means, it is important to block aerosol-borne template which can otherwise easily get into stock solutions and non-intended reactions. Always practice safe pipetting.

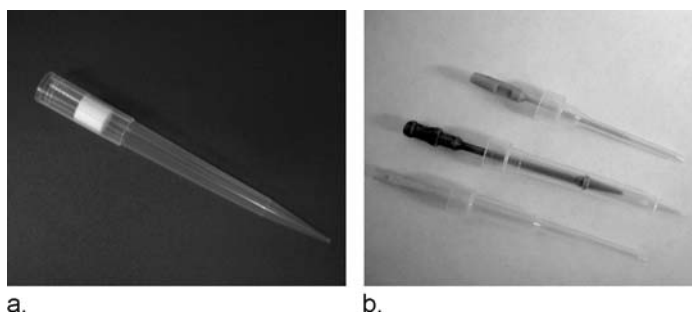


Figure 18.3 Micropipette tips for PCR. (a) Standard barrier tip (Rainin), designed to prevent aerosols that form within the tip from contaminating the micropipettor. (b) Positive displacement tips (Tri-Continent Scientific). The microsyringe style of these tips precludes aerosol formation altogether. Regardless of the method used, minimizing aerosol formation is imperative. Otherwise, the contamination of a micropipettor will result in amplicon chaos. Barrier tips are also quite useful for preventing radioactive contamination of micropipettors when working with isotopes as well for tissue culture applications in which aseptic technique is required.

Uracil-N-glycosylase

The enzyme uracil-N-glycosylase (UNG) is quite useful for the inhibition of contaminating carryover template. PCR reactions are assembled with dUTP, which functions as a dTTP analog, so that the amplified PCR product contains uracil. The inclusion of dUTP in the reaction does not influence the specificity of efficiency of product amplification. Prior to the next amplification, samples are incubated with the UNG enzyme which removes the base uridine wherever it is located. Consequently, the DNA backbone is susceptible to alkaline hydrolysis at the abasic sites created by UNG; breakage of any carryover uracil-containing template from previous reactions prevents reamplification. Fresh, thymine-containing template added to the reaction mix by the investigator is not a substrate for the UNG enzyme, and is therefore unaffected by this incubation. The UNG itself is subsequently destroyed by heating.

Primer design

Successful PCR requires a pair of properly designed primers⁶. One common source of confusion pertains to the nomenclature used to distinguish the two primers. For the sake of clarity, the various terms given to describe and distinguish each primer are presented in Table 18.3. For example, the *downstream primer* is that which base-pairs closest to the 3' end of the mRNA template, thereby supporting the synthesis of first-strand cDNA. This primer is also

⁶It is possible, in certain instances, to perform the reaction with a single primer species, such as when the same 5' to 3' sequence flanks the ends of the template.

Table 18.3 Essential Primer Nomenclature

5' primer	3' primer
Upstream primer	Downstream primer
Forward primer	Reverse primer
Sense primer	Antisense primer
(+) Primer	(-) Primer
Watson	Crick

Terms in the left-hand column are used interchangeably to describe the primer associated with the 5' end of the mRNA. Similarly, all of the terms in the right-hand column are used interchangeably to describe the primer responsible for amplifying sequences found closer to the 3' end of the mRNA. It is also correct to refer to primers more formally as oligonucleotides or just "oligos."

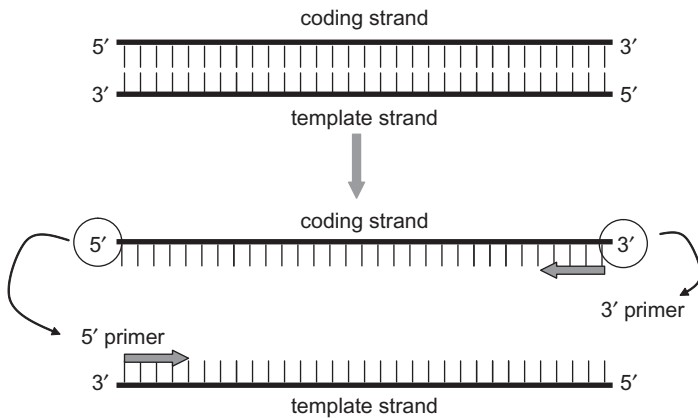


Figure 18.4 Essential primer nomenclature. The 5' primer base-pairs closest to the 5' end of the coding strand, while the 3' primer base-pairs closer to the 3' end of the coding strand.

referred to as the 3' primer, among other terms. The *upstream primer* base-pairs to newly synthesized first-strand cDNA sequences closer to the 5' end of the original mRNA (Fig. 18.4). This primer is also referred to as the 5' primer, among other terms. Most often, primers are designed by direct scrutiny of a published DNA sequence, be it genomic DNA or cDNA. By convention, only one strand of the DNA, the coding strand, written 5'→3', is published. When designing primers from sequences appearing in the literature, the reader is expected to know that there is indeed another strand base-paired and antiparallel to that which is published. The two most important rules for primer design, which pertain to how to extract the correct sequences in the correct orientation from the literature, are:

1. The upstream primer will have the same sequence as the published coding strand sequence.
2. The downstream primer sequence will be the inverted complement of the published coding strand sequence.

To ensure that the correct primer sequences have been selected, it is useful for the novice to write out and label both strands of the published sequence and ensure that the primers base-pair in an antiparallel configuration to opposite strands and that their respective 3'-OH groups point toward each other, thereby framing the sequence intended for amplification. One should not use software for this purpose until one has a clear understand of how all the molecules involved interact with each other. Any nucleic acid sequence can be amplified through the use of a properly designed primer pair under suitable reaction conditions, and primers that base-pair perfectly to the template are commonly known as full match primers.

Similar rules apply in the design of primers for RT-PCR: if an mRNA sequence is known, recall that it is the same sequence as the coding strand of the gene from which it was transcribed, except for the inclusion of uridine, rather than thymidine, in the transcript. Thus, the downstream gene-specific primer for RT-PCR, if oligo(dT) is not selected, should be the inverted complement of a region of the RNA upstream from the 5' end of the poly(A) tail. The upstream primer will be the same sequence as the targeted 5' region of the mRNA, except for the presence of thymidine, rather than uridine, in the primer. The other very important convention is that all primer sequences are always written 5'→3'. This applies equally to the upstream- and downstream primers. If a primer sequence is written incorrectly on an order form, the resulting primer, which won't work, will be entirely the fault of the person who filled out the form.

Example: An mRNA is known to have the sequence:

GAGCCGCACA	CCUGGGAUCC	GAUAAGAACA	CGGAGCGAAU	AUUUUUACUA
GCUAUGCUIA	GAGAGAUGAC	CGAGUAGCGA	CUAUUUUACC	AGCUUUUUUG
CGCGAUCGAC	CGAUCGACGA	AACAACAACA	CCCGCAACAG	CACGCGAGCG
AGACGAGAUU	CUACUAUGGG	UAAGUAGAAC	CCCAGACGAU	CACUCGUAA
CGCUAUCGAG	CUAGUAGGAU	GCGUGUGCGC	UCGUGCUAG	CGCUCGAGGC
UAAGUAGAAC	CCCAGACGAU	AGAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA
AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA
AAAAAAAAAA	AAAAAA			

Design a pair of primers, each 20 bases long that could generate a PCR product, after first-strand cDNA synthesis, spanning nucleotides 6–190.

Answer:

The sequence of the 5' primer: 5' GCA CAC CTG GGA TCC GAT AA

The sequence of the 3' primer: 5' ATC GTC TGG GGT TCT ACT TA

It is quite correct to list the nucleotides in groups of three so as to make reading the sequence easier.

In addition to sequencing data, oligonucleotide primers can be designed based on knowledge of the nucleotide sequence of a related gene in either the same or a different species, or knowledge of the amino acid sequence of the protein encoded by the gene of interest. The problem with the latter approach is the degeneracy of the genetic code: some amino acids have multiple codons.

When designing primers based on peptide sequence information, one strategy is to examine the peptide for clusters of amino acids with only one or two possible codons. The single-codon amino acids are methionine (Met, M) and tryptophan (TRP, W). One must realize, however, that the occurrence of five or six tandem single-codon amino acids is quite rare, and genes so endowed have, in all probability, already been cloned! The other amino acids all have two or more possible codons (Table 18.4). In the case of multiple codons, the investigator may have knowledge of preferential usage of one codon over another in a particular species, and this information can be acquired from a codon utilization table (e.g., www.kazusa.or.jp/codon). Without direct nucleic acid sequence information, the likelihood of designing a definitive oligonucleotide for use as a primer or as a hybridization probe is slim. As a general guiding principle, the six-fold degenerate codons (for arginine, leucine, and serine) should be avoided; if necessary, several 5' and 3' primers can be synthesized and then tested in PCR reactions in various combinations.

A degenerate primer is a mixture of primers having nearly identical sequences, with specific permutations at certain locations. Ergo, the molecules differ in at least one base position. Degenerate primers are useful when there are uncertainties pertaining to the expected homology with the template or when attempting to characterize sequence conservation across evolutionary time. The degenerate primer strategy facilitates annealing to and amplification of a variety of related sequences, some of which may vary significantly, depending on the exact cycling profile. For example, the primer

5' GGN CCG TCR TCR AAW GTC ARG TA

is a 23-mer that exhibits 64-fold degeneracy because of the presence of nucleotide variability at positions 3, 9, 12, 15, and 20. The overall degeneracy of the primer is determined by multiply together the fold-degeneracy observed at each position (Table 18.5). This means that the “primer” described here actually consists of 64 different types of molecules, corresponding to each of the 64 possible base permutations. Because certain genes show highly conserved regions across a breadth of organisms, degenerate primers can be helpful for fishing out a specific sequence from a genome. Now for the bad news: as primer degeneracy increases, primer specificity decreases. At least, however, the sequence being sought will not get way! One way to reduce primer degeneracy is to incorporate inosine, which is a nucleoside containing the purine base hypoxanthine, into the primer. Inosine can base-pair with adenine, cytosine, and guanine and, with slightly less efficiency, with thymine. As such, inosine is particularly useful for primers designed to span known single nucleotide polymorphisms (SNPs). Obviously, the inclusion of more than one inosine at positions where uncertainty exists in the design of a primer can substantially reduce the fold-degeneracy. Working with degenerate primers can also be useful when an amino acid sequence has been identified and one needs to work backward to design primers to isolate the corresponding gene. For the purpose of ordering primers, the standard symbols that are used to describe degenerate primers are listed in Table 18.5.

Table 18.4 Codon Utilization Table for Degenerate Primer Design

Amino Acid	DNA Codons	RNA Codons	Abbreviation	Fold Degeneracy	Symbol
Methionine	ATG	AUG	Met	1	M
Tryptophan	TGG	UGG	Trp	1	W
Asparagine	AAT AAC	AAU AAC	Asn	2	N
Aspartic acid	GAT GAC	GAU GAC	Asp	2	D
Cysteine	TGT TGC	UGU UGC	Cys	2	C
Glutamine	CAA CAG	CAA CAG	Gln	2	Q
Glutamic acid	GAA GAG	GAA GAG	Glu	2	E
Histidine	CAT CAC	CAU CAC	His	2	H
Lysine	AAA AAG	AAA AAG	Lys	2	K
Phenylalanine	TTT TTC	UUU UUC	Phe	2	F
Tyrosine	TAT TAC	UAU UAC	Tyr	2	Y
Isoleucine	ATA ATC ATT	AUA AUC AUU	Ile	3	I
Valine	GTA GTC GTG GTT	GUA GUC GUG GUU	Val	4	V
Proline	CCA CCC CCG CCT	CCA CCC CCG CCU	Pro	4	P
Threonine	ACA ACC ACG ACT	ACA ACC ACG ACU	Thr	4	T
Alanine	GCA GCC GCG GCT	GCA GCC GCG GCU	Ala	4	A
Glycine	GGA GGC GGG GGT	GGA GGC GGG GGU	Gly	4	G
Arginine	AGA AGG CGA CGC CGG CGT	AGA AGG CGA CGC CGG CGU	Arg	6	R
Leucine	CTA CTC CTG CTT TTA TTG	CUA CUC CUG CUU UUA UUG	Leu	6	L
Serine	TCA TCC TCG TCT AGT AGC	UCA UCC UCG UCU AGU AGC	Ser	6	S
Nonsense codons					
	TAA	UAA (Ochre)	Stop		
	TAG	UAG (Amber)	Stop		
	TGA	UGA (Umber)	Stop		

Basic rules

The following are some of the characteristics of useful primers for PCR-associated applications. Be aware, however, that no matter how much brain power goes into primer design, vindication comes only when the primers actually work!

Table 18.5 Standard IUB Base Abbreviations for Primer Design

Base	Common symbol(s)	Abbreviation	Fold degeneracy
Adenine	A	A	1
Cytosine	C	C	1
Guanine	G	G	1
Thymine	T	T	1
Uracil	U	U	1
Inosine	I	I	1
Purine	A/G	R	2
Pyrimidine	C/T	Y	2
Adenine, Cytosine	A/C	M	2
Guanine, Thymine	G/T	K	2
Adenine, Thymine	A/T	W	2
Cytosine, Guanine	C/G	S	2
Cytosine, Guanine, Thymine	C/G/T	B	3
Adenine, Guanine, Thymine	A/G/T	D	3
Adenine, Cytosine, Thymine	A/C/T	H	3
Adenine, Cytosine, Guanine	A/C/G	V	3
Any base	A/C/G/T	N	4

The table contains the conventions for representing the specific order of bases, and degree of primer degeneracy, for the unambiguous synthesis of specific oligonucleotides needed to support PCR. Note also that inosine is actually a nucleoside that contains the base hypoxanthine. IUB = International Union of Biochemistry.

1. Ideally, primers should be 19–24 bases long in the region that base-pairs to the template. Longer primers may be used, though their thermodynamic behavior becomes less predictable as length increases. With shorter primers, one runs the great risk of promoting non-specific primer annealing, because as a primer becomes shorter, its discriminatory abilities diminish. To an extent, minimum primer length is species-specific. A good rule for determining minimum primer length whilst maintaining primer specificity is to solve for the exponent:

$$4^X > Y$$

where

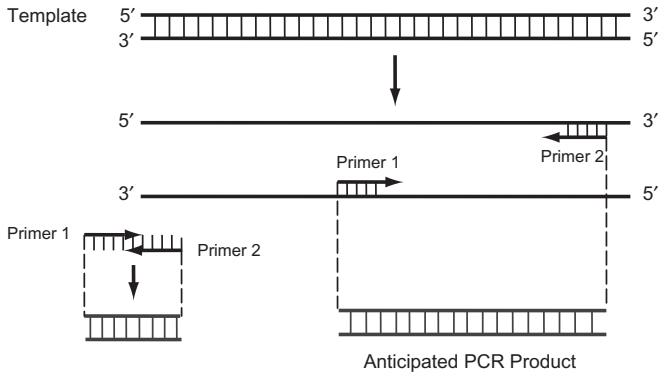
4 = the number of possible nucleotides that make up the primer

X = the length of the primer

Y = the genome size of the organism for which the primers are being designed.

For example, the diploid size of the human genome is about 6.6×10^9 bp (equivalent to about 6 pg DNA per cell). Thus, $4^X > 6.6 \times 10^9$. By solving for X, it is determined that, at least theoretically, a 17-base primer (a 17-mer) is long enough to find only the DNA sequence to which the primer is targeted, and not another unrelated sequence due to the random occurrence of bases. After all, 6.6×10^9 base pairs is a lot of sequence. In practice, a 17-mer remains in what should be thought of as the gray area, where a primer may or may not impart the necessary level of sensitivity. Therefore, a bit longer is better.

2. In a PCR reaction, it is far more important to match primers by T_m than by length or by sequence. T_m differences of more than a few degrees, especially among shorter primers, are likely to have a profound negative effect on the outcome of the reaction in terms of specificity and yield.
3. Mismatches at the 3' end are fatal. Mismatches (locations where a base in the primer does not hydrogen bond with the template) that involve the 3'-most nucleotide, and often the penultimate nucleotide, will not support primer extension. Polymerases require the 3'-most nucleotide carrying the requisite 3'-OH to be perfectly base-paired to the template. When primers are designed by trial and error, there is often no way of knowing if the 3' end is perfectly base-paired to the template. If such an ambiguity exists, it may be wise to target a different region of the gene; often moving just a few bases 5' or 3' from the original primer target site may be all that is necessary.
4. The 5' end of a primer can have rather exotic overhangs when the investigator wishes to generate a PCR product with new sequences on one or both ends. Useful sequences to add include any combination of DNA sequencing promoters, RNA transcription promoters, restriction enzyme sites, and sequences that allow the PCR product to be recognized and amplified by a completely different primer pair for downstream applications.
5. Each primer sequence should be examined to ensure that it does not exhibit the potential for intramolecular hybridization due to the presence of inverted repeats; if present, the primer could base-pair to itself. The formation of what is best thought of as a primer hairpin causes a marked reduction in free primer concentration, thereby drastically reducing the product yield. This is more of a problem than commonly recognized. Because of the extremely rapid hybridization kinetics associated with oligonucleotides, it should be intuitive that a primer molecule will base-pair to the spatially closest sequence to which it is complementary. If the 5' end of a primer is complementary to its own 3' end, then hybridization is sure to occur and the magnitude of the problem is dependent on the annealing temperature. Moreover, intramolecular base pairing can promote non-specific base pairing to the correct template, though in the wrong location, due to the presence of a few non-hybridized bases at the 5'- and 3' ends of the primer, resulting in a variant of an "insertion by PCR". Happily, the Internet is rife with free software tools that can be used to analyze primer sequences for potential difficulties. One of the best ways to survey the variety of gratis primer characterization tools is to go to any search engine and type in "primer analyzer" or some related variant of search terms. Many of the websites will calculate T_m , molecular weight, propensity for hairpin formation, and likelihood of interaction between primers, to name a few useful applications. Interestingly, analyzing the same primer sequence at different websites does not always produce identical results. The only real test of primer design is how the primer(s) perform in the thermal cycler.
6. Each primer pair should be examined to ensure that the primers are not complementary to each other near their 3' ends. When primers base-pair to each other, the synthesis of primer-dimer is favored, detracting greatly from the synthesis of the intended PCR product. The incidence of primer-dimer is greatest when the primers are GC-rich at their 3' ends. If primer scrutiny suggests that primer-dimer may occur, changing the length of one or both primers by as few as one or two nucleotides will usually rectify this problem.



7. Primers should be designed with average GC content. In humans, the GC content of the genome is about 41%. In this laboratory, primers with 40–55% G+C are considered average. This parameter is important because guanine and cytosine base-pair with three hydrogen bonds between them, while adenine and thymine base-pair with only two. This will preclude one primer from base-pairing to the template more tenaciously than the other primer.
8. A particular primer should be designed with an average distribution of nucleotides, as opposed to having a disproportionate density of guanine and cytosine nucleotides at one end of the primer, and with adenine and thymine at the other. This will preclude having one end of the primer base-pair with and denature from the template with a significantly different efficiency than the other end.
9. Repeated motifs should be avoided. For example 5' ATGATGATGATGATGATG may cause specificity problems.
10. There should be no more than two Gs or Cs among the last five bases at the 3' end of a primer. For example, using the primers 5' TAGCTAGCCTAATATATCTGCGC and 5' TCATCCAATATCGATCGTAGCGC would strongly favor primer-dimer formation with themselves and with each other. This would be disastrous.
11. PCR can often be improved by designing a primer(s) with a single G or C at the 3' end. Some authors argue in favor of the absolute necessity of using PCR primers with GC dinucleotide at the 3' end. The logic behind this argument is that a GC-rich clamp⁷ of sorts (GG, GC, CC, or CG) will increase the efficiency of priming by stabilizing the oligonucleotide on the template. While this approach may be helpful in cases where a primer shows relaxed specificity, it may also cause primer-dimer formation depending on the calculated annealing temperature and the base sequence in the remainder of the oligonucleotide. Most of the established high-precision primers used in this lab do not end with one of the GC dinucleotide permutations.
12. A single gene-specific primer may be all that is necessary to support PCR when one end of the template has been ligated to an adapter sequence or otherwise tailed. For example, the downstream primer imparts template specificity while the upstream primer anneals to the adapter. This approach is characteristic of RACE PCR, described below, and the template-switching strategy described in Chapter 17.

⁷The term “GC-clamp” is more commonly used to describe a string of guanine and cytosine bases appended to the 3' end of primers used for specialized techniques such a denaturing gradient gel electrophoresis. The term is appropriate in this context as well.

T_m considerations

It is clear that the design of each primer is of fundamental importance to the success of PCR; of equal importance is how the primers behave together in a reaction tube. Designing useful primers necessitates promoting a proper balance between template specificity, thermodynamic stability when base-paired to the template, and capacity of one primer to function with the other(s) to support RT-PCR. The probable collaborative behavior of one or more pairs of oligonucleotide primers is best described in terms of the T_m of each primer involved. The T_m is that temperature at which 50% of the possible annealing events between primer and template have occurred and 50% have yet to occur; thus the T_m is best thought of as an equilibrium temperature at which one primer base-pairs to the template for every primer that melts or dissociates from the template. As noted above, it is important to match primers closely by T_m .

Estimating T_m

T_m is a direct function of the length of the primer, the base sequence of the primer, and the concentration of salt (Na^+) in the reaction. One rapid method for determining T_m is to assign 2° for each adenine and thymine and 4° for each cytosine and guanine:

$$T_m = 2^\circ(\text{A}+\text{T}) + 4^\circ(\text{G}+\text{C})$$

This method is reliable for 11-mer to approximately 28-mer primers. When the T_m is determined to be greater than 68°, this calculation is no longer reliable. This method is very widely used to estimate the probable T_m for each of a pair of primers. Conveniently, when primers are synthesized commercially, the specification sheet that accompanies them generally includes the T_m calculation, most often for 50mM NaCl. When designing primers, it is best to match them closely based on T_m rather than by length. As suggested previously, primer sequences should be analyzed *in silico* before they are actually synthesized, in order to predict their probable behavior.

The T_m defines an annealing equilibrium point; it should be fairly obvious though, that one does not wish to saturate *only* 50% of all complementary target sites, but rather all of them. Thus, knowledge of the T_m of each primer allows the calculation of the temperature of annealing (T_a), the second step of each PCR cycle. To promote the greatest specificity of primer annealing, one may wish to use an annealing temperature of only a few degrees below the T_m , for example, $T_a = T_m - 5^\circ$. If annealing 5° below the T_m fails to generate product, one should consider repeating the experiment at $T_m - 10^\circ$, and so forth. Incrementally decreasing the annealing temperature is most often required when the one or more mismatches exist between primer and template. While mismatches will be tolerated at lower temperature, be acutely aware that by lowering the annealing temperature, non-specific priming is also favored.

Precision T_m calculations

The estimation of thermodynamic behavior based on inventorying the number of each base, as described above, is usually a fairly accurate predictor. For greater precision, especially in cases where multiple primers are to be synthesized and used in various combinations, it is best to use an on-line primer analyzer that will not only calculate thermodynamic behavior but also propensity for the primer to participate in primer-dimer formation or the base-pair to itself. The other nice thing about primer analysis websites is that the value for monovalent- and divalent cation concentrations can be changed. This gives the investigator some flexibility in PCR itself as well as the potential to use either or both of the primers as hybridization probes for library screening, Southern analysis, and the like.

On-line resources

Everyone is fully aware the internet revolutionized the age of information. Many biotech vendors, and certainly all vendors that manufacture custom oligonucleotides, have some type of primer design or analysis software at their website. These tools are remarkably handy not only for basic primer interaction analysis and T_m determination, but also for the amount of related information that is calculated instantly. For example, when a primer sequence is input using an on-line oligonucleotide analyzer, the software computes optical density (OD), molecular weight (MW), the number of nanomoles per OD (nmol/OD), the number of micrograms per OD ($\mu\text{g}/\text{OD}$), extinction coefficient, GC-content, the sequence of the complementary strand to which the primer is expected to base-pair, and other related properties. That this information is available immediately is quite convenient and would be needed later on in any event when performing concentration determinations and when going back and forth between mass- and molar concentrations.

On a related, recall that degenerate primers actually consist of two or more different primer species, depending on the degree of degeneracy. For example the 20-mer sequence GCTARATGCTATCTACBATC actually consists of a mixture of the molecules:

```
GCTAAATGCTATCTACCATC
GCTAAATGCTATCTACGATC
GCTAAATGCTATCTACTATC
GCTAGATGCTATCTACCATC
GCTAGATGCTATCTACGATC
GCTAGATGCTATCTACTATC
```

because of the degeneracy at nt 5 (R = A, G) and at nt 17 (B = C, G, T). In instances such as this, an on-line primer analyzer will also provide the minimum, maximum, and mean T_m for the primer pool.

Multiplex primer design

“Multiplexing” is the process of using several primers simultaneously to amplify one or more target sequences, though the primers need not be degenerate. It is

important, however, to maintain fairly tight control over the melting temperature range of the primers involved. One worthwhile strategy is to prepare separate pools of primers, the members of each of which exhibit a similar T_m . Of course, it is also possible to use only one primer, a “best-guess” oligonucleotide sequence (also known as a “guess-mer”), at a lower annealing temperature. While relaxed stringency may favor semi-non-specific hybridization, the important point is that the desired sequence is amplified, albeit among false positives, which can be eliminated in the later stages of sequence identification. At lower stringency, not all of the bases of a primer will necessarily base-pair with the template. Mismatches, individual bases that do not hydrogen-bond to the template, are expected to abound at lower temperatures, though the primer may continue to anneal enough to support elongation. If mismatches are suspected it may be wise to perform the first 2–3 PCR cycles 5–10° below the desired annealing temperature. This strategy will stabilize the primer and permit primer extension along the template in the presence of multiple mismatches, which is critical for successful PCR. Lastly, remember that if a pool of primers is used, necessitated by target sequence ambiguity, the effective concentration of the correct primer sequence will be much lower compared to a homogeneous primer.

Optimization procedures

The polymerase chain reaction has become as commonplace in the molecular biology laboratory as a set a micropipettors. Without a doubt, it has revolutionized the way many questions pertaining to cell biology and other disciplines are approached. Having a high confidence level in a PCR-based assay comes from reproducibility which, in turn, is highly dependent on optimization of the reaction. Optimization procedures are frequently required whenever a new set of primers is utilized.

There are numerous potential difficulties, some obvious, others not so obvious, that the investigator must address. Several of these are listed below, and are best taken to heart. Succinctly, the efficiency and specificity of PCR are influenced by *each* component of the reaction.

1. The standard method for sample preparation involves the preparation of a large volume of master mix which contains all but one of the reaction components. This master mix is then aliquoted into individual reaction tubes, followed by the one variable component that was missing from the master mix. Master mix preparation minimizes tube to tube variations attributable to pipetting technique. Also be sure to prepare slightly more master mix than the anticipated need. For example, if four samples are to be tested, multiply the per-tube volumes of each component by 4.5 or by 5.0. This will prevent any master mix shortage as each sample tube is prepared.

A typical PCR reaction consists of 10mM Tris-Cl, pH 8.3 (at 20°); 1.5mM $MgCl_2$; 50mM KCl; 200 μ M dNTP; 0.5 μ M of each oligonucleotide primer; 2.5 U *Taq* polymerase/100 μ l reaction; 10^2 – 10^5 copies (<1 μ g) per 100 μ l reaction. Each of

these components is subject to change, though the parameters listed here represent a good place to begin.

2. Pay particular attention to the final magnesium concentration when using a 10× PCR reaction buffer without MgCl_2 , requiring addition by the investigator. This divalent cation binds the dNTPs and the template (negatively charged phosphodiester backbone) quantitatively, thereby exerting a measurable effect on the efficiency and specificity of the reaction.
3. Be sure to ascertain template quality. PCR amplifies segments of DNA that lie between the primers. Thus, the reaction is a bit more tolerant of partially degraded DNA and RNA than some of the classical techniques in molecular biology. Templates that are severely degraded are unlikely to support the amplification of anything. Wherever possible, always check template quality by minigel electrophoresis before proceeding to PCR.
4. The final concentration of each primer needed to support PCR is typically 0.1–1.0 μM , beyond which product yield drops off precipitously. If newly designed primers fail to generate a product when annealing close to the T_m , then reduce the annealing temperature incrementally, and do so at least twice. If the primers continue to fail to produce PCR product, don't waste time with primers that don't work: select alternative primer sequences. Primers can be ordered at any scale of synthesis, and may be shipped in lyophilized form. Just prior to first use primers should be reconstituted in sterile H_2O or modified TE buffer (10 mM Tris, pH 8.0; 0.1 mM EDTA). Next, the concentration of each primer should be determined by UV absorption. *Never* assume that the mass of the primer reported on the side of the tube or on the product spec sheet is completely accurate. An empirical determination is absolutely critical to ensure a symmetric reaction. Failure to use equimolar quantities of the upstream and downstream primers is a sure way to favor spurious product synthesis.
5. Primer purity is an important parameter of amplification efficiency. It should be intuitive that highly purified primers support high reaction efficiency and specificity. Primer purification means that, at a minimum, residual compounds associated with primer synthesis are removed. It is also important to understand that full length primer molecules are invariably contaminated with truncated molecules (also known as failure sequences), the presence of which could negatively impact the outcome of an experiment. There are several levels of purification to which primers can be subjected by the manufacturer. While the cost of some purification methods is included in the purchase price of the primer, the more elaborate purification methods will incur an additional charge. Typical purification options include desalting, cartridge purification, reverse-phase HPLC, PAGE, and gel filtration.
 - a. Desalting is usually included in the price of the primer and is generally more than satisfactory for most mainstream PCR applications, up to 35-mers. The purpose of desalting is to remove the residual chemistries and unnecessary organic materials associated with the synthesis of the primer, removal from the column, and primer deprotection⁸. For primers longer than 35 bases, or when very pure primer preparations are required, more extensive cleanup procedures will be helpful.
 - b. Cartridge purification is a rapid method that produces primer purity (80–90% pure) often approaching that which is possible with HPLC. The method is based

⁸Deprotection is the process by which the chemical groups protecting the reactive base and phosphate components of the nucleotides are removed by hydrolysis with ammonium hydroxide.

on differences in hydrophobicity due to the presence of a 5' DMT (dimethoxytrityl) group in full-length primers and the absence of this chemical group in non-full-length molecules, i.e., those molecules that were incomplete at the end of the synthesis process. Cartridge purification is useful for primers up to 50 bases in length. The overall efficiency of quantitative assays, cloning, and site-directed mutagenesis reactions can be improved through the use of cartridge purification or HPLC.

- c. Reverse-phase high performance liquid chromatography (HPLC) often generates primer purity exceeding 95%. The purification principle is similar to cartridge purification. HPLC can be used to purify greater amounts of a primer, meaning more than 1 μ mol, but is not recommended for primers greater than 50 bases. HPLC-purified primers often require an extra day or two before they are shipped from the manufacturer.
 - d. Polyacrylamide gel electrophoresis (PAGE) purification is the ultimate purification tool, with primer purity routinely as high as 99%. This is the method of choice for the purification of primers between 50–100 bases because stiff polyacrylamide gels offer resolving power that can separate molecules that differ by as little as a single base. Investigators should expect a reduced yield of PAGE-purified primers due to the required recovery of the primer from the gel. This method of purification also requires an extra day or two prior to shipping.
 - e. Gel filtration is recommended for oligonucleotides intended for use *in vivo*, in conjunction with cell culture, or for antisense assays because of the potential for cytotoxicity in the presence of any residual synthesis of purification chemistries. Gel filtration can be used as a stand-alone method of primer purification or in conjunction with the other methods described above. As a stand-alone method, gel filtration will remove the inorganic salts, residual reaction chemistries, and very short truncated sequences, after which the primer can be lyophilized, resuspended in a suitable buffer, and quantitated.
6. Test the components of each new PCR kit before using it with precious template. Many PCR kits contain a test template and control primers, used to show that the reagents are functional. Failure to generate a product can just as easily be due to compromised reagents as to poor primer design or degraded template.
 7. Optimize PCR conditions and perform subsequent reactions in the same physical thermal cycler, and not just another instrument with the same model number. Moreover, due to enormous variation in efficiency among different models, reactions thought to be optimized in the older machines suddenly manifest new PCR products in the newer machines. These “new” bands were, in fact, probably being generated in the lesser efficiency machines, but their relatively low mass on the gel put them below the level of detection. If reactions suddenly stop working altogether, check to make sure that someone has not accidentally altered the thermal cycler programming: modification of an optimized cycling profile usually is not a good thing.
 8. Each component of the cycling profile (denaturation, annealing, and extension) should be kept as short as possible with the exception of certain very specific applications such as long-range PCR. Short cycles, along with the rapid movement between temperatures (ramping) that is commonplace with later-model thermal cyclers, tend to generate highly reproducible reactions. In cases where the annealing temperature (T_a) is $>65^\circ$ it may be helpful or even necessary to use a two-temperature cycling profile, rather than the usual three-cycle approach. For example, the annealing

and extension components can be combined into a single step that is performed between 68° and 72°.The thermal cycler would be programmed as follows:

Initial melt	94°	2 min	
Denature	94°	15 s	} 30 cycles
Anneal/extend	94°	15 s	
Soak	4°	∞	

This is possible because longer probes with higher T_m will associate with the template more rapidly than shorter primers at higher temperatures. Moreover, all thermostable enzymes commonly used for PCR have excellent polymerase activity between 68° to 72°.

9. Consider investing in an “optimization kit” which often consists of combinations of chemistries known to enhance the PCR process. For example, such kit might contain several 10× buffer formulations with varying concentrations of magnesium and separate aliquots of various adjuncts to enhance the robustness of the reaction. Known PCR adjuncts include *E. coli* single-strand binding protein (Chou, 1992), formamide (Sarkar *et al.*, 1990), ammonium sulfate (Olive *et al.*, 1989), bovine serum albumin (BSA) (Paabo *et al.*, 1988), dimethyl sulfoxide (Cheng *et al.*, 1994; Winship, 1989), betaine⁹ (Rees *et al.*, 1993; Henke *et al.*, 1997) and glycerol (Cheng *et al.*, 1994). The influence of each of these adjuncts is summarized in Table 18.6. One worthwhile strategy is to select the best 10× PCR buffer, based on product yield and absence of non-specific product, and then to add the various adjuncts one at a time. This will allow the investigator to identify the proper magnesium concentration and adjunct combination needed to optimize the reaction for a particular template and primer set. This will also save time and money whenever a new set of reaction conditions is used.
10. Beware of the following known inhibitors of PCR: humic acid, polysaccharides, hemoglobin, IgG, lactoferrin, high bacterial concentrations, urea, phenol, chloroform, guanidinium thiocyanate, excess NaCl, sodium dodecyl sulfate (SDS), N-laurylsarcosine (sarcosyl), ethanol, formalin, bromophenol blue, and EDTA, to name but a few.
11. Avoid mineral oil as a reaction overlay. Not only is mineral oil messy but its use also correlates nicely with diminished reaction robustness and non-reproducibility of data. If you do not have thermal cycler with a heated lid in your lab, then make friends with someone who does.

Thermostable polymerases

Many of the historical difficulties associated with the synthesis of representative cDNA carry over into the realm of RT-PCR. Amplification efficiency and overall sensitivity can be improved by using combinations of thermostable polymerases. These so-called enzyme blends were originally formulated to support long-range PCR, a technology which commonly results in the synthesis of genomic-length PCR products, i.e., greater than 4kb. In this laboratory the long range technology is used to support the efficient amplification of cDNAs that correspond to

⁹Betaine = N,N,N-trimethylglycine.

Table 18.6 Effects of Adjuncts that Enhance PCR

PCR adjunct	Working concentration	Effect
Ammonium sulfate	15–30 mM	Influences denaturation and annealing profiles of the template and primers by increasing overall ionic strength
Betaine	0.8–1.0 M	Reduces secondary structure in GC-rich regions; increases yield and specificity of PCR products
Bovine serum albumin (BSA) (non-acetylated)	10–100 µg/ml	Stabilizes <i>Taq</i> (and other enzymes, too); particularly useful for ancient DNA or templates containing PCR inhibitors, especially humic acid
Dimethyl sulfoxide (DMSO)	1.0–10.0%	Decreases incidence of template depurination; also reduces secondary structure; good for GC-rich templates
Formamide	1.0–5.0%	Increases stringency of primer annealing, especially in GC-rich template domains
Glycerol	5–10%	Increases <i>Taq</i> stability; lowers template T_m and reduces secondary structure
Non-ionic detergents (NP-40, Tween 20, Triton X-100)	0.1–1.0%	Stabilizes <i>Taq</i> polymerase; can neutralize carryover SDS from DNA extraction procedures
SSBP (single-stranded binding protein)	0.7–1.5 µg/assay	Accelerates primer annealing; inhibits extension of mismatched primers

exquisitely low abundance mRNAs as well as longer mRNAs in the cytoplasm, and much longer unspliced hnRNA. This amplification strategy provides greater sensitivity and, overall, is quite beneficial. A typical enzyme blend, for example, might consist of *Taq* polymerase with another thermostable polymerase, such as *Pwo*. *Taq* is an efficient (processive) enzyme with respect to its intrinsic 5'→3' polymerase activity, though the half-life of this enzyme is only about 40 min at 95°, and it lacks 3'→5' exonuclease activity (proofreading). Thus, if *Taq*, or any other non-proofreading enzyme, makes an error, the error becomes a permanent part of the PCR product. Other thermostable enzymes, such as those from the genus *Pyrococcus* exhibit greatly enhanced thermostable qualities and also exhibit a proofreading capability. The combination of *Taq* and one of these other enzymes, used in conjunction with a modified buffering system, is much more efficient and amplifies with higher fidelity than is possible with either enzyme alone. Moreover, in instances when amplification is extended beyond 30 cycles, PCR product can continue to accumulate, because the second enzyme more or less takes over after the *Taq* has been, quite literally, burned out. The thermostability of some of the commonly used PCR enzymes is presented in Table 18.7. Various enzyme blends are readily available from most vendors.

Table 18.7 Stability of Selected Polymerases Used for PCR

Enzyme	Organism	Half-Life	Proofreads?
Taq	<i>Thermus aquaticus</i>	40 min at 95°	No
Tfl	<i>Thermus flavus</i>	40 min at 95°	No
Tth	<i>Thermus thermophilus</i>	20 min at 95°	No
Tli	<i>Thermococcus litoralis</i>	400 min at 100°	Yes
Pfu	<i>Pyrococcus furiosus</i>	12 h at 95°	Yes
Pwo	<i>Pyrococcus woesei</i>	2 h at 100°	Yes
Deep Vent _R TM	<i>Pyrococcus GB-D</i>	23 h at 95°	Yes
Vent _R [®]	<i>Thermococcus litoralis</i>	6.7 h at 95°	Yes
Stoffel Fragment	<i>Thermus aquaticus</i>	80 min at 95°	No

Positive controls

The number of positive controls should be limited because they can easily contaminate stock solutions and cause unmitigated havoc in various downstream applications. As needed, positive controls should not be the same size as the experimental amplicon. That way, if positive control template does manage to contaminate stock solutions or other components of the reaction at any level, it will be easily distinguishable as a separate band on a gel or by melting-curve analysis in the case of real-time PCR. Avoid working with highly concentrated positive control template stock solutions and be sure to assemble the positive control reaction last.

Negative controls

Each set of reactions and the corresponding gel should contain at least one negative control, the nature of which is dependent on the experiments being performed. Examples of negative controls include a tube into which no template was added (NTC; no template control), a which contains RNA template that was not exposed to reverse transcriptase (RT⁻ control; tests for the presence of genomic DNA), and two tubes that contain only one of the two primers (tests for self-priming; need to perform this only once with each new primer pair). In this laboratory, negative control reactions are assembled from the master mix after all of the tubes containing experimental template have been assembled, but before assembly of the positive control reaction tube.

Hot start PCR

A common problem that results in spurious product formation is associated with the assembly of reactions at room temperature or on ice in a non-hot-start format: during the initial ramping (raising the reaction temperature) to 94° needed to achieve initial melting and then to initiate the cycling, the temperature of the sample obviously must pass 72° on the way up to 94° (Fig. 18.5).

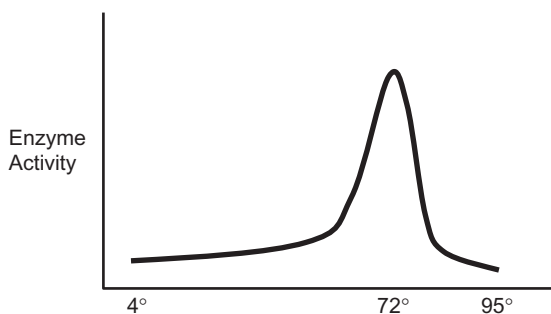


Figure 18.5 Spurious product formation can result prior to the initial melt in a non-hot start PCR format because of the burst of polymerase activity that is observed as the reaction temperature passes 72° on the way up to 94°. Hot start approaches inhibit polymerase until after the reaction components have reached denaturation temperature for the first time.

As 72° is the optimal temperature for primer elongation (at least for *Taq*), non-specific PCR product can be generated even if the primers are imperfectly base-paired to sites on the template where they would not have base-paired under more stringent conditions. This is especially true of primer-dimer formation. Any non-specific product made at this stage is more than likely to be coamplified with the intended template throughout the reaction. To preclude this difficulty, hot start PCR is specifically designed to keep the enzyme inactive until the temperature of the reaction is at least higher than the melting temperature of both primers. In so doing, the incidence of mispriming is drastically reduced if not eliminated altogether, which would otherwise lead to spurious PCR product formation.

Among the numerous strategies used to prevent the synthesis of any product at all until both the primers and the template have been completely denatured prior to the onset of the first cycle is the use of an anti-*Taq* antibody which inhibits the polymerase activity of the enzyme until the antibody is heat denatured, such as the TaqStart™ antibody (Clontech) and JumpStart™ Taq (Sigma). Alternatively, enzyme activity can be suppressed by associating it with an aptamer, i.e., a short nucleic acid molecule with a protein-binding ability (GeneAmp® AccuRT RNA PCR kit; Applied Biosystems). Other hot-start approaches (some fashionable, others out of style) include (1) manual hot start, in which one reaction component is withheld from the tube until the samples have been pre-heated; (2) preheating the thermal cycler to 95° and then placing the sample tubes directly into a hot block in order to reduce the ramping time; (3) the use of AmpliWax® products (Applied Biosystems) to separate the components of the reaction until the first heating step melts the wax barrier, allowing all reaction components to mix so amplification can begin; (4) sequestering the requisite magnesium in a wax bead which, upon melting, releases a precise quantity of magnesium so that the polymerase becomes active; and (5) using

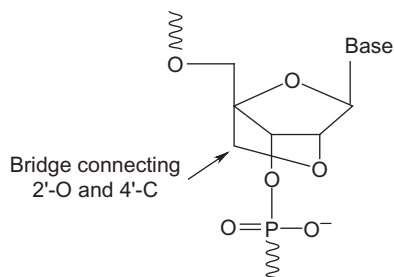


Figure 18.6 Locked nucleic acids (LNAs) are modified nucleotides in which the 2' oxygen and 4' carbon atoms are joined by an extra bridge. Oligonucleotides that include LNAs show enhanced specificity, sensitivity, and hybridization stability.

modified reaction chemistry by which *Taq* is active only when heated to high temperatures, and loses activity as the temperature is reduced. The application of hot start technology is an important facet of any RT-PCR based assay and can be the deciding success factor in assays featuring “difficult templates”.

Locked nucleic acids

Locked nucleic acids (LNAs; Koshkin *et al.*, 1998) are modified RNA nucleotides in which the 2'-O and 4'-C atoms of the ribose are joined through a methylene bridge (Fig. 18.6). This additional bridge limits the flexibility normally associated with the ring, essentially locking the structure into a rigid conformation. These nucleic acid analogs are also referred to in some circles as “inaccessible RNA”.

LNAs can be inserted into both RNA and DNA oligonucleotides, and their presence increases the T_m by 2 to 4° per LNA. The inflexible nature of these molecules greatly enhances hybridization stability. Further, oligonucleotides containing LNAs offer tremendous discriminatory power, allowing these molecules to distinguish between exact match and mismatched complementary target sequences with very little difficulty. LNAs are also widely respected for their enhanced biostability in blood and other biological fluids. LNA oligonucleotides are well-suited for antisense experiments, many PCR procedures, ribozyme design, *in situ* hybridization, SNP (single nucleotide polymorphism) identification, microarray screening, and aptamer functionality, to name but a few possible applications. One of the nice things about LNA oligonucleotides is that the same platforms used for DNA synthesis can be used for LNA synthesis, meaning that labs with (or access to) a DNA synthesizer may begin making LNA oligonucleotides immediately.

When an oligonucleotide contains standard nucleotides and LNA nucleotides in the same molecule, the term “mixmer” is sometimes invoked to more aptly describe their nature. When describing the sequence of a mixmer, one of the conventions is to use lower case letters for standard nucleotides, and upper case letters for LNA nucleotides, while another convention is to place a + symbol

in front of upper case letters to represent an LNA nucleotide. For example, the primer gctGcacgTcatcgatcATCtcatgc, a 26-mer with LNAs in positions 4, 9, 18, 19, and 20, can also be written as GCT+GCACG+TCATCGATC+A+T+CTCATGC. By comparison, the non-LNA version of this primer is written as gctgcacgtcatcgatcctcatgc or GCTGCACGTCATCGATCATCTCATGC. When designing LNA-containing oligonucleotides, one should follow the basic rules of primer design described above, and one should pay particular attention to the location and number of LNAs. For example, a typical 18-mer should contain a maximum of 7–8 LNAs. Also, try to avoid stretches of more than 4–5 consecutive LNAs, which would result in very tight hybridization in that region. Stretches of LNAs are to be avoided close to the 3' end of an oligonucleotide. Finally, be sure to match the T_m of the primers, as usual, keeping in mind that each substitution of a standard nucleotide with an LNA increases the T_m .

Touchdown PCR

Problems with RT-PCR specificity? Make sure that a hot start formulation is being used. Also, using a touch down PCR cycling profile can drastically improve the overall specificity of the reaction. Succinctly, touchdown PCR involves programming the thermal cycler such that the initial annealing temperature is higher than the T_m of any primer in the reaction tube. With each passing cycle, the annealing temperature is reduced by a predetermined increment, usually 0.5° , until the annealing temperature is reduced to about 5° less than the lowest T_m among the primers in the tube. As a result, virtually no non-specific product is generated in the first few cycles, and when the temperature is low enough so that primer annealing does occur, the primers will only interact with exact match sequence. By the time the reaction reaches its lowest annealing temperature, the expected product is in a large excess over non-specific product that may have formed in the last few cycles when the annealing temperature was below the T_m . This approach is renowned for imparting great specificity to reactions involving “difficult” primers (Fig. 18.7) and does a remarkable job of eliminating spurious product. Many of the newer thermal cyclers contain a touchdown format, and the user enters the starting and ending annealing temperatures which, of course, will be unique to the primer set.

Internal controls

Inconsistencies in technique and day-to-day variation among investigators are often responsible for the introduction of error at nearly every level of scientific inquiry, beginning with one's own laboratory skills all the way through to data analysis. To address at least one aspect of this potential problem in transcription assays, a commonly used approach is the measurement of an internal control or endogenous standard, the expression of which is *not* expected to change as a function of experimental manipulation. These are often referred to

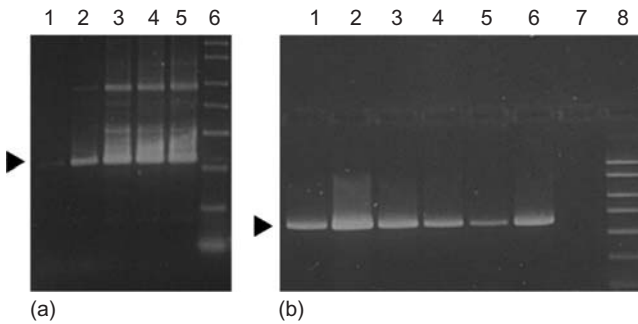


Figure 18.7 Touchdown PCR. Primers will occasionally interact with template in a non-specific manner, even at temperatures close to the T_m . By gradually reducing the annealing temperature from starting temperature *above* the T_m , only perfectly matched primer:template duplexes will be tolerated and amplified. (a), Five first-strand cDNA samples (lanes 1–5) were subjected to PCR with primer annealing at 57° . Ideally, these primers should have produced a single, 365 bp PCR product, indicated by the arrow; other non-specific product is clearly visible. Lane 6 is the PCR molecular weight standard. (b), Six first strand cDNA samples (lanes 1–6) to which a touch down PCR profile was applied. The annealing temperature was reduced from 65° to 55° at a rate of 0.5° per cycle. At 55° , 15 additional cycles were performed. A single product is observed at the predicted location in lanes 1–6. Lane 7 is empty and lane 8 contains the PCR molecular weight marker. The PCR products in panels (a) and (b) are identical in size, but were not electrophoresed on the same gel.

as reference- or housekeeping genes, the expression of which is associated with basic (basal) cellular metabolism. Housekeeping genes, characteristically, are expressed constitutively, meaning that transcripts are produced continuously from those gene loci. The general idea is identify at least one gene, the expression pattern of which does not change, or at least shows minimal modulation, as a function of experimental manipulation.

The transcripts from the housekeeping gene allow normalization (meaningful comparison) among two or more samples, thereby facilitating calibration of gene expression. The housekeeping gene may be assayed either before or after the assay of the experimental gene(s) on the same blot or in the same sample series. In this regard, it is worth reiterating the fact that Northern analysis is plagued with a general lack of sensitivity due to the mechanics of the technique. Data of much greater sensitivity are achievable via nuclease protection analysis (Chapter 15) and the very popular quantitative RT-PCR methods (Chapter 19). Regardless of the method used, assay of a housekeeping gene is essential. In addition, some investigators choose to spike an RNA sample with a known quantity of an exogenous reference template (discussed below) for precision measurement of both reverse transcription efficiency and the overall mass of PCR product at the end of the reaction. Be aware, however, that while the spiking approach has merit with respect to accurate quantification of the expression of specific transcripts, it does not help in determining the overall transcriptional health of the cell.

Investigators should be aware that currently there is no all-purpose internal or “reference” transcript applicable to every system and every situation. This is a fact of life. Many of the commonly assayed, constitutively expressed housekeeping genes are, alas, subject to variation by the cell. This is especially true in the assay of RNA associated with the disease phenotype. Happily, one is usually able to identify a gene or two for each model system without too much difficulty, and it is wise to assay at least two different genes so that ensuing conclusions are accurate and convincing. PCR methods have demonstrated that these genes, once thought to be relatively invariant, can change dramatically in terms of transcript abundance in response to experimental challenge, requiring the suitability of each to be determined empirically. Finally, it is usually a good idea to pick a housekeeping gene with a different physiological function than the genes being assayed for variation. For example, if cells are treated with a compound that inhibits microfilament formation, one should pick as an internal control a gene that is unrelated to the cytoskeleton; in this case tubulin would not be the best choice.

Below are some of the commonly reported reference genes. While this list is by no means intended to be exhaustive, it does highlight those genes whose transcripts are most frequently reported as controls for the purpose of reliability of results. The appropriateness of a particular transcript for use as an internal control transcript is entirely dependent on the biochemistry of the system under investigation. While there is no gene that will remain invariant under all conditions, genes that show the least variation under a variety of experimental conditions are those best suited as internal control genes.

1. β -actin, a constitutively expressed transcript of medium abundance: its isoforms are found in nearly all cell types (Kinoshita *et al.*, 1992). Historically the first transcript to be used as a standard, the use of β -actin as a control transcript remains commonplace. Numerous reports, however, have demonstrated repeatedly the variation of this gene in response to a variety of external stimuli and cytoplasmic cues (Leof *et al.*, 1986; de Leeuw *et al.*, 1989; Dodge *et al.*, 1990; Solomon *et al.*, 1991; Spanakis, 1993).
2. GAPDH (glyceraldehyde-3-phosphate dehydrogenase), manifested as a transcript of medium abundance, encodes a key enzyme required for glycolysis. The expression of the GAPDH transcript as a reference gene or control is in widespread use, including this lab. As is the case with β -actin, GAPDH gene expression is also subject to modulation under a variety of experimental conditions (Alexander *et al.*, 1990; Mansur *et al.*, 1993; Bhatia *et al.*, 1994; Lau *et al.*, 2000).
3. Cyclophilin, a protein involved in the isomerization of peptide bonds, is a highly conserved gene, the RNA from which is of medium abundance in the cytoplasm. Cyclophilin A is also known as peptidylprolyl isomerase A (PPIA). While commonly assayed as a control transcript, variation in cyclophilin mRNA has been demonstrated (Danielson *et al.*, 1988; Haendler and Hofer, 1995; Feroze-Merzoug, *et al.*, 2002).
4. The histones are a family of proteins needed for nucleosome formation and regulation of transcription. The transcripts are abundant and are often assayed as housekeeping genes (Farrell and Greene, 1993; Robert, *et al.*, 2002).
5. Hypoxanthine-guanosine phosphoribosyltransferase (HPRT, *also* HGPRT), an enzyme well known for its role in the salvage pathway, whereby nucleotides are recycled for use by the cell. HPRT has been shown to be constitutively expressed at

rather low levels (Pernas-Alonso *et al.*, 1999). However, HPRT levels are known to fluctuate and are typically elevated in proliferating cells (Steen *et al.*, 1990).

6. Dihydrofolate reductase (DHFR), an enzyme necessary for the synthesis of purines, thymidylate, and glycine. As such, the expression of this gene is absolutely necessary for cell growth. DHFR is has been cited widely as a housekeeping- or reference gene (Linton *et al.*, 1989; Horikoshi *et al.*, 1992; Oleary *et al.*, 1994; Balajee *et al.*, 1999).
7. 26S proteasome, consisting of several subunits, is responsible for proteolysis in eukaryotic cells (ubiquitin-proteasome pathway). Primers and probes targeting specific subunits are often used as controls in various transcription assays (Bassett *et al.*, 2005; reviewed by Voges *et al.*, 1999).
8. Other genes, including fibronectin (abundant transcript; extracellular matrix), transferrin receptor (medium-low abundance; iron uptake) ubiquitin (small protein that targets other proteins for degradation via proteasomes), lamin (nuclear intermediate filament proteins), tubulin (microtubule formation), and porphobilinogen deaminase (PBGD; cytosolic enzyme in heme biosynthesis; Chretien *et al.*, 1988) mRNAs are widely reported as control transcripts. The protein product of each of these genes has a unique role in the cell. In general, these transcripts are less likely to experience disturbing fluctuations in mRNA levels, although there are always exceptions to this generalization. Test these first!

The word on transcription controls

A common source of confusion is the relationship between transcription controls and loading controls in the discernment of the overall transcriptional response of the cell to experimental challenge. mRNA and its unspliced precursor hnRNA are the products of the RNA polymerase II activity in eukaryotic cells. Thus, when studying changes in mRNA abundance, it makes the greatest sense to use a transcription product of RNA polymerase II as an internal control, particularly since the eukaryotic RNA polymerase show differing sensitivities to various compounds, of which α -amanitin (Chapters 1, 16) is a notorious example. It is the opinion of this Author that the products of RNA polymerases I and III, rRNA and tRNA, which are described below, are better suited as loading controls, meaning that they are used to show that an equivalent mass of RNA is present in each sample. A well-designed set of experiments should consist of transcription controls (mRNA) and loading controls (usually rRNA).

1. 28S and 18S rRNA, highly abundant posttranscriptionally-regulated products of the ribosomal genes, are often cited as nearly invariant transcripts of great predictability as an internal control (de Leeuw, 1989; Bonini and Hofmann, 1991). These transcripts are an essential component of mature ribosomes. It is not difficult to imagine a situation where an experimental manipulation causes a large overall reduction in mRNA synthesis all the while rRNA and tRNA are being produced, or at least sustained in the cytoplasm, at near control levels. The reverse is also possible.
2. 16S mitochondrial rRNA, moderately abundant transcripts that are localized in the mitochondria, is utilized as an internal control with lesser frequency than its cytoplasmic counterparts. The expression of these mitochondrial genes, purportedly, is influenced far less by experimental manipulation than are genes localized in the

nucleus (Tepper *et al.*, 1992). However, the fact that these are ribosomal genes, rather than mRNA, should incite the same concerns expressed in item 1, above.

3. tRNA consists of a number of high abundance transcripts which shuttle amino acids to the ribosomes during translation. The total mass of tRNA is occasionally measured, using a tRNA-specific probe, as a loading control. tRNA is transcribed by RNA polymerase III, and is therefore better suited as a loading control than as a gene expression control, for the same reasons cited in item 1.

Analysis of PCR products

There are two common methods by which PCR products are analyzed. The first method, end-point analysis, is performed using the time-honored method of electrophoresis at the conclusion of the amplification. The other method, known as real-time PCR, allows the investigator to collect data pertaining to PCR product accumulation at the end of each cycle. Inasmuch as real-time PCR is described in detail in Chapter 19, the following comments pertain to routine end-point analysis.

At the conclusion of a typical reaction, usually consisting of 25–30 cycles, the PCR tube is opened and an aliquot of each sample is run out on a gel along with an appropriate MW standard. In most settings, agarose gels (1–3%) represent the matrix of choice for the evaluation of PCR products and are able to resolve almost any size PCR product. These gels may be prepared with or without ethidium bromide added to the gel solution after it has been melted and cooled to about 55°. Alternatively, gels may be stained with SYBR Green or SYBR Gold *after* electrophoresis, for superior sensitivity.

A variety of agaroses are available from most vendors. One may elect to use standard- or low melting temperature agarose, or an agarose grade that favors separation of low molecular PCR products. In this lab, SeaKem GTG® (Lonza) is routinely prepared as a 2.5 to 3.0% solution in 1× TAE. This recipe favors exceptional resolution of PCR products ranging from 200 to 1000 bp. Be sure to use caution when melting high agarose concentrations such as these in a microwave, as the tendency for the agarose to boil over is quite high, both in the microwave and immediately after heating. To reduce the boil-over potential, let agarose sit in TAE buffer for about 5 to 10 minutes before heating in order to allow the agarose beads to swell. Then heat the agarose to boiling while continuing to exercise caution. Great care must be taken when preparing these high concentration gels to prevent injury to the investigator and to avoid burning large clumps of agarose in the flask or beaker, especially when using a hot plate.

The intensity of a band on a gel is associated with its abundance at the end of the reaction, and a reasonable estimate of the mass on the gel can be made if it is electrophoresed along with a mass standard, which consists of well characterized dsDNA molecules in which the concentration of each band has been precisely measured. A better method for determining reaction yield, however, is to clean up the reaction using a silica-based PCR purification column and then measuring the purified PCR product spectrophotometrically.

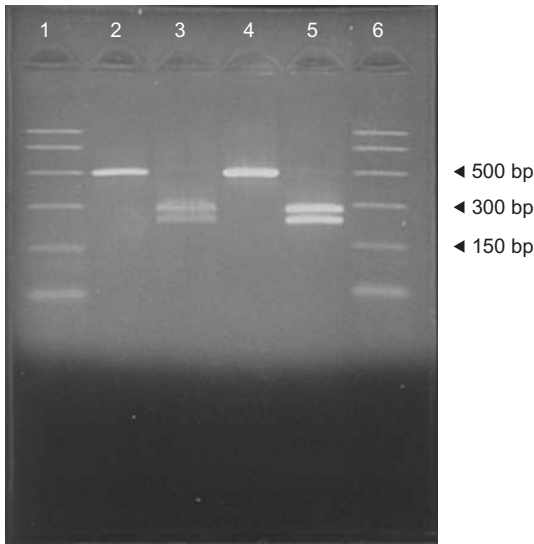


Figure 18.8 The presence of a known restriction enzyme site in routinely produced PCR products can be used for putative product identification. The 500 bp PCR product in lanes 2 and 4 has a single *Mbo*I site which, upon digestion, renders the expected 227 and 273 bp fragments (lanes 3 and 5). This precludes the need for constant sequencing and is a good spot check of the specificity of the assay. If the PCR product did not cut, then an immediate examination of all aspects the amplification would be performed. Lanes 1, 6: PCR mw standard.

Another concern that is often expressed when PCR products are electrophoresed pertains to the identity of the product itself. In real-time PCR, product identity, and the presence of spurious products, is determined by melting curve analysis: a dsDNA product of a particular length and base composition should denature at a characteristic temperature. This analysis is performed in the real-time thermal cycler immediately after the last cycle. Of course, DNA sequencing is the ultimate identification tool, though this is not practical for every product generated in every reaction. When end-point analysis is performed in this lab, many of the products that are expected are known to contain what we refer to as a diagnostic restriction site that can be used to (putatively) identify a product (Fig 18.8). Every so often an aliquot of a PCR product is digested with the restriction enzyme that is expected to cut it. If the appropriate banding pattern is observed, this is taken as an indicator of the identity of the product without the need for sequencing. This approach is very worthwhile.

RT-PCR quality control points

Among the more notorious difficulties encountered when preparing a nucleic acid sample for PCR is the integrity of the template material, a facet that takes on an entirely new meaning when working with RNA. For those laboratories

actively investigating gene expression, it is likely that some PCR variant¹⁰ is being used to amplify one or another transcript with primers and parameters that have long since been optimized. The frustration index rises when one is unable to generate a product from a freshly prepared sample of RNA, or from an RNA sample that “used to work”. The following should be pondered, along with the optimization items discussed previously.

1. Assess the quality of the RNA just prior to use by electrophoresis. The minigel format is more than adequate to simply show the presence of the 18S and 28S rRNAs. RNA that has been stored in aqueous buffer, rather than as an ethanol precipitate, degrades rapidly. This is exacerbated when the sample has been subjected to repeated freeze–thaw cycles. In this lab, previously purified RNA is heated to 65° for 5 min and then electrophoresed in a 1.2% agarose minigel in 1× TAE (no denaturants needed; no blotting); the sole purpose is to assess the integrity of the sample. As long as there is reasonable definition to the 28S and 18S rRNAs, the investigator may proceed confidently.
2. Test each pair of primers without template. This is necessary to show that the stock solutions are not contaminated with an amplicon as well as to give any indication of primer-dimer. This is the so-called “no template control.”
3. Test each primer alone, with template. This is the only true way to show that self-priming is not a possibility and that the primer pair is specific. Fig. 18.9 shows the surprising results of precisely this type of QC experiment in an evaluation of a new set of primers.
4. Test each sample of RNA by omitting the reverse transcriptase from the first-strand synthesis reaction. Then use the contents of this tube as template material for PCR. In

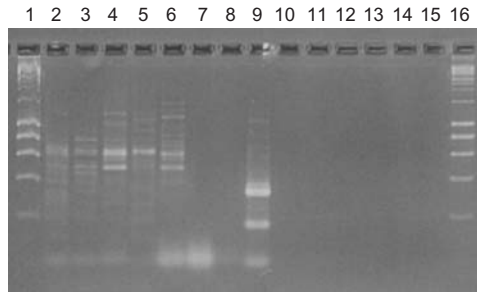


Figure 18.9 Testing for primer self-priming. Two primers designed for use in a transcription assay were tested independently for propensity for self-priming. Lanes 2–6: 100 ng first-strand cDNA from five different samples was added to aliquots of PCR master mix that contained only the 5′ primer. Lanes 10–14: 100 ng first-strand cDNA from the same six samples was added to aliquots of master mix that contained only the 3′ primer. Lanes 7 and 15 were negative controls (no primer), Lane 8 was empty, lane 9 contained a positive control, and lanes 1 and 16 contain PCR molecular weight standard. The numerous, unexpected variety of PCR products generated with only the 5′ primer clearly demonstrated its capacity for self-priming. Upon observing these data, the 5′ primer for the assay was redesigned.

¹⁰Related RT-PCR techniques are described in detail in Chapters 19, 20 and 22.

this sample, known as the “no RT control” or the “RT-minus control”, no products should be observed. In the absence of reverse transcription of the RNA template, the only bands that could possibly be generated will have come from a contaminating genomic DNA in the sample, assuming that there is no amplicon contamination anywhere in the pipeline. The standard protocol in most labs performing any type of RT-PCR is to DNase the samples (Appendix F) at the time of isolation, after which the RT- control is performed.

5. Test difficult RNA samples with primers that may not be of direct experimental interest, but that are at least known to generate a PCR product each time RNA is isolated, intact and clean, from a particular biological source. This can be performed using very boring primers, such as those for β -actin, histone, or GAPDH just to show that one can amplify something (anything) from the RNA in question. If a PCR product is evident with control primers, then the reason for failure to generate a product from experimental primers can be investigated more intelligently. Succinctly, no signal can mean no transcription *or* that the RNA sample is poor and unable to support amplification of anything. This latter possibility will have been validated, or ruled out, through the use of these control primers.
6. Phantom bands may be asymmetric product. Upon electrophoresis, an unexpected band that appears to be about one-half the size of the expected product is often attributed to the preferential amplification of one strand. This occurs when one primer is present in a molar excess over the other primer, and this is a more frequent occurrence than one might think. This can be remedied easily by checking the primer concentrations spectrophotometrically, regardless of what is printed on the label on the tube. This will ensure that subsequent dilutions of primer stock solutions produce equimolar primer concentrations in future reactions. If this was the problem, the phantom bands will disappear.
7. Be judicious in the use of positive control templates, the products from which should be distinguishable electrophoretically, or by melting curve analysis, from the expected product. If a positive control amplicon manages to contaminate a stock solution, everything in the lab will come to a grinding halt.

Non-PCR methods for confirming PCR-derived data

PCR is an efficient amplification tool, but it is also a biased reaction. It is clear that some sequences amplify with greater efficiency than others and, as the number of cycles increases, the true abundance relationships among the starting material (mRNA \rightarrow cDNA) are distorted. One of the most convincing approaches for validating changes in gene expression from RT-PCR experiments is to use a non-PCR method to show that same trend. For example, demonstrating up- or down-regulation of a transcript by nuclease protection (Chapter 15) or by reverse Northern analysis (Fig. 18.10) is hard to contradict. There are other methods by which to confirm PCR-derived data, too. Because these methods vary widely in sensitivity, it might be helpful to review “Farrell’s Sensitivity Index” in Chapter 19.

It is also true that alternative methods may be difficult to employ because the transcript of interest may be of very low abundance and, therefore, below the level of detection by some less sensitive technique(s).

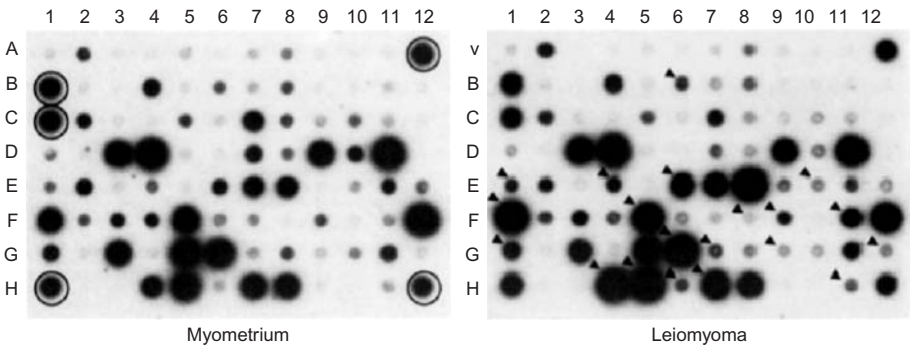


Figure 18.10 Reverse Northern analysis. Differentially expressed genes are detected by reverse Northern analysis using mRNAs as probe template. Two sets of blots are made with an equal quantity of PCR products. Positive controls (G3PDH) are marked with circles. Genes that have been up-regulated are indicated with arrows. Leiomyoma is a benign smooth muscle tumor and myometrium is normal smooth muscle tissue from the wall of the uterus. Reprinted with permission from Macmillan Publishers Ltd: *Cell Res.* 12, 215–221, copyright 2002.

Many of the methods described in this book are simple to perform while others are rather labor-intensive. The intention here is not to burden the investigator with additional tasks prior to submitting data for publication. Rather, the emphasis here is the importance of confirming one's findings using other methods, to the extent possible. This becomes particularly important when only subtle changes in gene expression are detected; given the sensitivity of PCR to all aspects of reaction composition and assembly, even minor variations in pipetting can have an amazing impact on the product abundance at the end of the amplification protocol. Moreover, because labs often have principal investigators, students, and/or postdocs working on various aspects of the same project, it is not uncommon for tubes to become mixed up (date, template ID, concentration), leading to conflicting or erroneous data. Just as one may drive to a destination via a number of different routes, one should also be able to draw the same conclusions via different methods.

This strategy goes a long way toward convincing colleagues, critics, competitors, and peer-reviewers.

Related techniques

5' RACE PCR

One of the key drawbacks of traditional cDNA synthesis is the failure to generate full-length cDNA molecules that capture sequences corresponding to the 5' end of the mRNA. This is due principally to historical difficulties with reverse transcriptase in the synthesis of first-strand cDNA, the mechanics of PCR, and up- and downstream primer design. The resulting loss of sequence makes it impossible to clone full-length sequences or map the transcription start site (TSS) at a particular

locus. A method known as RACE (rapid amplification of cDNA ends; Frohman *et al.*, 1988) was developed mainly for cloning sequences that characterize the 5' end of a transcript (5' RACE) or the 3' end (3' RACE), or for cloning full-length cDNA molecules when only partial sequence information is available.

There are two strategies for performing 5' RACE. The older method involves the synthesis of first-strand cDNA as usual, accomplished with a downstream primer, either oligo(dT) or, ideally, a gene-specific primer that base-pairs closer to the 5' end of the transcript. Recall that any deoxyoligonucleotide can provide the requisite 3'-OH needed to initiate reverse transcription. Next, the newly synthesized first-strand cDNA molecules are tailed, meaning that extra nucleotides of known sequence are added enzymatically to the 3' end of the cDNA by the somewhat unique enzyme terminal deoxynucleotidyl transferase. Known a bit more casually as "terminal transferase" this enzyme will add a few additional nucleotides in a non-template-dependent manner to either blunt- or protruding 3'-OH groups on either single- or double-stranded DNA, as shown in Fig. 18.11. The extra nucleotides added in this manner constitute what is known as the anchor sequence¹¹. The addition of one species of nucleotide is often referred to as homopolymeric tailing; thus a poly(G) anchor can anneal a poly(C) anchor (upstream) primer to support PCR. Similarly, a poly(A) anchor will be able to anneal a poly(T) anchor primer. Using this approach, amplification by PCR will only require one gene-specific (downstream) primer, not two. A major limitation of this method, however, is the fact that any first-strand cDNA molecules made by reverse transcriptase are candidates for the homopolymeric tailing, even truncated cDNAs that did not reach all the way to the 5' end of the mRNA template. For this reason, false positives are to be expected and, depending on the positioning of the downstream primer, may be present in large numbers. Even worse, the co-amplification of the truncated species is in effect a competition reaction which greatly detracts from the ability of the reaction to amplify the desired product, especially when it is present in low abundance. Finally, the homopolymeric nature of the anchor primer creates a substantial T_m imbalance when used with to the downstream gene-specific primer needed to support amplification.

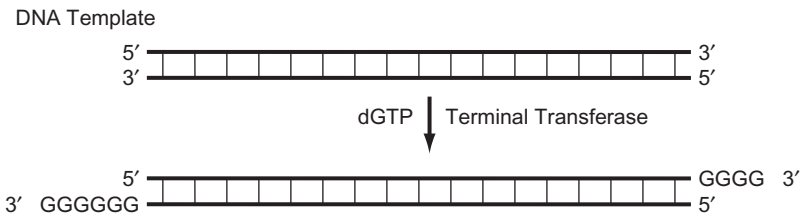


Figure 18.11 Action of terminal deoxynucleotidyl transferase. This unusual enzyme, also known as terminal transferase, is capable of adding nucleotides to blunt- and protruding 3' ends of DNA in a non-template-dependent manner. This activity can be exploited for a number of PCR and cloning applications.

¹¹In the past, this cloning method was also known as anchor PCR (Loh *et al.*, 1989) and one-sided PCR.

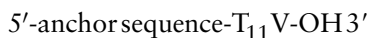
5' RLM-Race

A newer, excellent method for cloning the 5' end of transcripts is known as RNA ligase-mediated RACE (5' RLM RACE; Maruyama and Sugano, 1994; Shaefer, 1995). This technique drastically improved the ability to capture only cDNA molecules corresponding to 5' end of the mRNA, rather than truncated cDNAs or cDNAs that were primed from partially degraded transcripts. Succinctly, the method involves using the enzyme T4 RNA ligase to join a sequence-defined RNA oligomer to the 5' end of only full-length, undegraded mRNA. Upon reverse transcription, the ligated RNA oligomer becomes the 3'-most end of the cDNA to which the upstream primer will anneal.

The RLM-RACE process is outlined in Fig. 18.12. After running a gel to ensure that the RNA sample is very high quality, RLM-RACE begins by incubating the RNA in the presence of calf intestinal phosphatase (CIP) to remove any exposed 5' phosphate groups. The 5' cap structure is resistant to CIP, so the mRNAs that will be dephosphorylated are those from which the cap has already been removed or those molecules which are already partially degraded. The result is that damaged, non-full-length transcripts now have a 5'-OH structure. Next, the enzyme tobacco acid pyrophosphatase (TAP) is used to remove the 7^mG nucleoside along with the α and β phosphates. This exposes a single phosphate group at the 5' end of only full-length mRNAs, a necessary substrate for T4 RNA ligase. An RNA anchor (adapter) sequence is then ligated to the decapped mRNAs. All of the other molecules that were partially degraded or already decapped at the onset of the experiment are not candidates for ligation because those molecules now have a 5'-OH group, and not the required 5'-PO₄. The RNA to which the anchor was ligated is now ready for reverse transcription. While it is true that all RNA molecules that are able to anneal the downstream primer will be reverse transcribed, the only first-strand cDNAs that will be amplified are those which contain the adapter sequence and can anneal the upstream anchor-specific primer. The RLM-RACE approach has the added advantage of using nested anchor primers that conform more closely to the standard rules of primer design with respect to base distribution and average GC content: anchor primers direct toward homopolymerically tailed cDNAs are either 100% GC-rich or 0%! In a recent investigation in this lab, RLM-RACE was used to identify multiple TSSs from one gene locus and in a tissue-specific manner (Bassett, *et al.*, 2004).

3' RACE PCR

Cloning the sequences at the 3' end of the transcript (3' RACE) does not pose the numerous challenges associated with 5' RACE. The standard 3' RACE method involves the priming mRNAs first-strand cDNA synthesis using a downstream primer with the generic structure:



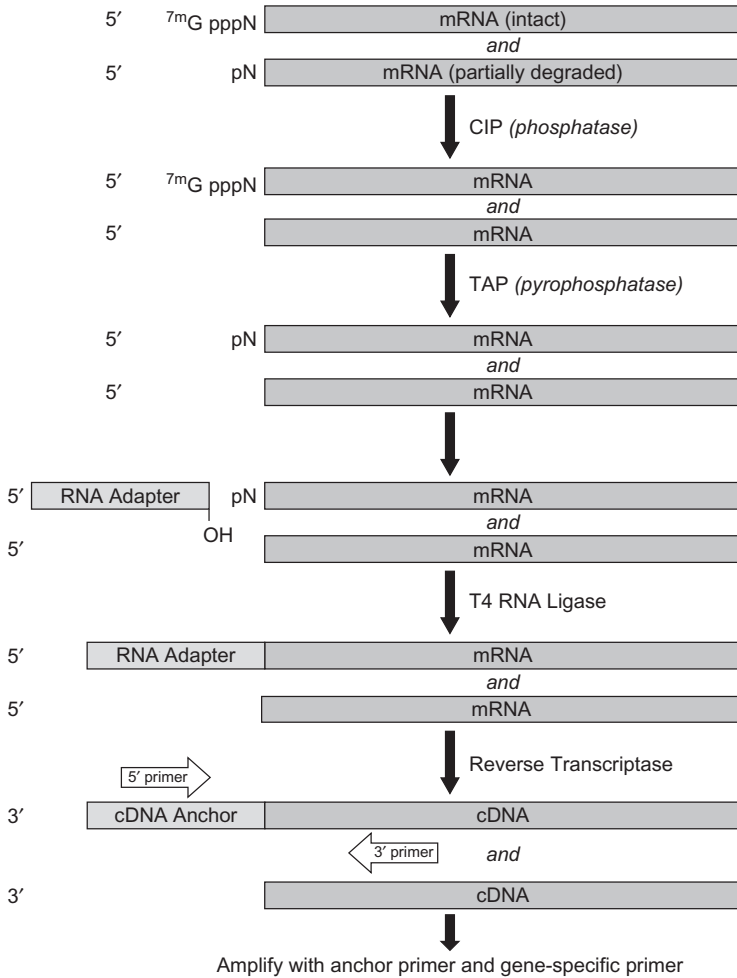


Figure 18.12 RNA ligase-mediated RACE. The mechanics of this assay favor the cloning of only intact mRNA molecules that manifest a 5' cap. The novelty of this approach is the ligation of an RNA adapter after calf intestinal phosphatase (CIP)-mediated dephosphorylation of partially degraded mRNA. Then, incubation of full-length mRNA with tobacco acid pyrophosphate (TAP) removes the cap, thereby exposing the requisite 5' for the ligation event. Non-mRNA, corrupted mRNA, and first-strand cDNA molecules that failed to reach the RNA adapter during the ensuing reverse transcription are non-amplifiable. RLM-RACE is an extremely effective tool for transcription start site (TSS) identification.

where

1. The anchor is a unique sequence located 5' to an oligo(dT) tract;
2. T₁₁ is the minimum length of the oligo(dT) tract that should be used for stability reasons;

3. The 3' end of the primer (V) is a single nucleotide anchor consisting of adenine, cytosine, or guanine (*see* Table 18.5). This ensures that the primer positions itself at the 5'-most end of the poly(A) tail so that reverse transcription proceeds as far into the coding region of the mRNA as possible. It is also possible to design the downstream primer as 5'-anchor sequence-T₁₁VN-OH 3' with a dinucleotide anchor at the 3' end for added stability. Following cDNA synthesis, PCR amplification of cDNAs is supported through the use of an anchor (downstream) primer that is complementary to the anchor, and the use of a gene-specific (upstream) primer (Fig. 18.13). While one might envision 3' RACE by simply using oligo(dT) as the downstream primer, experience strongly advises against the use of any AT-rich sequence: oligo(dT) is 100%!

Nested PCR

Occasionally, the failure to generate a PCR product may be related to the exquisitely low abundance of the template or, perhaps more commonly, to the purity of the template. In one strategy an aliquot of a completed PCR reaction, which manifests no product, is diluted 1:100 into a fresh reaction mix and subjected to an additional 30 cycles. However, a technique involving two sets of primers, known as nested PCR, is a widely regarded supplemental amplification strategy,

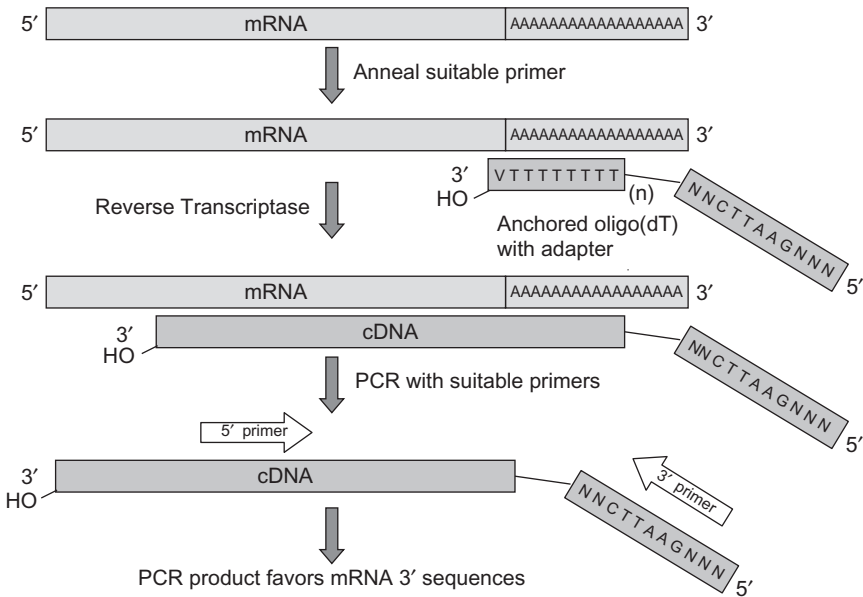


Figure 18.13 3' RACE. The approach is used when the goal is to amplify sequences that correspond to the 3' end of the original mRNA. This method is far less complicated than either traditional 5' RACE or RLM-RACE. The use of an anchored oligo(dT) primer with an adapter at the 5' end positions the primer at the 5' end (the beginning) of the poly(A) tail, and provides a novel sequence at the 5' end of the first-strand cDNA to which a 3' primer can be annealed (following synthesis of second-strand cDNA).

compared to simply running through more cycles; this is especially true when there is a purity issue to be resolved.

In the nested PCR approach, one of the two sets of primers (the inner primer set) is designed to base pair completely within the region amplified by the other set of primers (the outer primer set). The relative positioning of these primers is shown in Fig. 18.14. Succinctly, the sample is amplified using the outer primers for 30 cycles, after which an aliquot (dilution optional) is amplified for an additional 15–30 cycles using the inner primers. If properly designed, the outcome of the nested PCR approach is usually most gratifying. Upon electrophoresis, the product generated with the nested primers (secondary product) migrates a bit more rapidly than the product(s) produced with the outer primers (primary product), because the secondary product is shorter by at least the combined size of the outer primers. One must be certain to examine all four of the proposed primer sequences to ensure that they conform to the standard rules for primer design, as discussed above.

Long-range PCR

The concept of long-range PCR applies to the synthesis of genomic-length PCR products which are widely accepted to be greater than 4000 bp. Almost unthinkable in the early days of PCR, improvements in component purity, modifications of the reaction chemistry, and the clever selection of thermostable polymerases facilitate the routine amplification of very large pieces of DNA (Fig. 18.15).

The very essence of the long-range PCR process is the innovative use of two thermostable enzymes in the reaction tube, rather than one. These so-called enzyme blends often consist of *Taq* polymerase, which is highly processive but does not proof-read, and another thermostable enzyme that

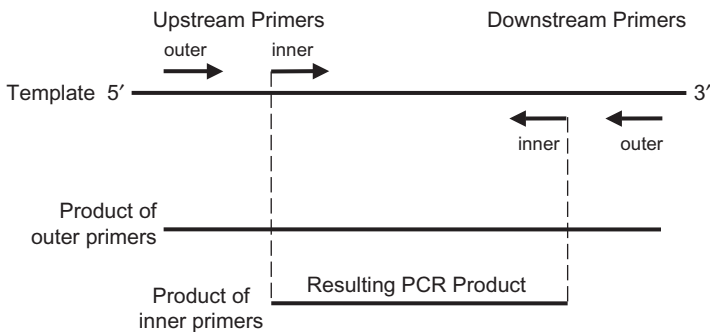


Figure 18.14 Nested primers. Two sets of gene-specific primers are often utilized when issues of product purity surround a PCR-based assay. A primary PCR amplification is performed using the outer the primers. This is followed by a second PCR amplification using the inner primers, designed to lie completely with the region recognized by the outer primers, and a diluted aliquot of the product(s) of the primary reaction. This approach confers tremendous specificity to the assay by minimizing the amplification of any non-specific product that may have tainted the primary reaction.

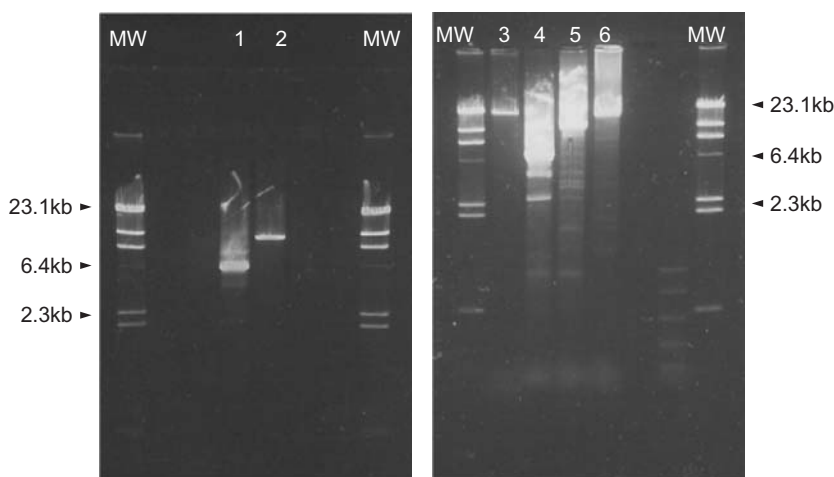


Figure 18.15 Long-range PCR. The products on this gel range from 2.3 to 23.0kb. Note the location of the λ -*Hind* III mw standard in the outside lanes on the gels. This procedure is very likely to require optimization to eliminate smearing (truncated product formation) between the major bands observed upon electrophoresis. Lane 1, 4.8kb; lane 2, 9.3kb; lane 3, 23.0kb; lane 4, major product is 9.4kb; lane 5, 15.0kb; lane 6, 23.0 kb.

usually has a longer half-life, is less processive, but does proofread. Inasmuch as *Taq* polymerase makes mistakes, the two-enzyme format can produce very long PCR products with high fidelity.

There are other unique long-range PCR parameters and long-range reactions are far more likely to require optimization than standard reactions.

1. Primers used for long-range applications tend to have a T_m in 60–68° range and, as always, the T_m should be similar for both.
2. Long-range reaction buffers are more alkaline (pH >8.8). This improvement minimizes the risk of depurination of the template during cycling. Since the upstream and downstream primers often anneal to the template thousands of base pairs way from each other, it is imperative that the template be intact to support amplification.
3. Long-range PCR is far less tolerant of partially damaged template than more traditional forms of the reaction; therefore, template integrity and purity are much more important because of the rather unforgiving nature of the long-range reaction format. Template material should be checked electrophoretically (size) and spectrophotometrically (concentration and purity) before use.
4. Long-reaction reactions are often supplemented with one of the PCR adjuncts described in Table 18.6 which lowers the T_m of the template so that denaturation can be performed at 92° degrees rather than 95°. Further, primer extension can be reduced from 72° to 68°. While these temperature reductions may at first seem trivial, they go a long way toward extending the life of the polymerase once cycling is underway.
5. The $MgCl_2$ concentration is more likely to require titration, and is often higher (1.75–2.25 mM) than is associated with standard PCR applications.

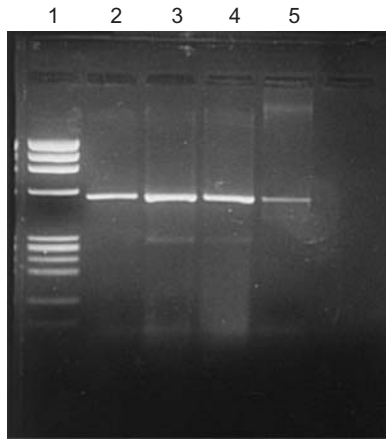


Figure 18.16 Enhanced PCR amplification using enzyme blends. 50 ng of RNA was reverse-transcribed, purified, and then aliquoted into four different PCR formulations. After 25 cycles, reaction products were analyzed on a 3% agarose gel containing 0.5 $\mu\text{g/ml}$ ethidium bromide. Lane 1: ΦX174 mw standard. Lanes 2–4: three different enzyme blends (*Taq* with another thermostable, proof-reading polymerase). Lane 5: *Taq* polymerase alone. Digital image analysis revealed at least a 2-fold increase in the mass of the resulting 580 bp product when an enzyme blend was used, compared to *Taq* alone. (Farrell (1999). *Biotech International* 11, 1304. With permission.)

6. The concentration of nucleotides (dNTPs) is usually increased to 300–500 μM for long-range PCR. Some investigators favor the use of ultrapure nucleotides while others find the cocktail from a standard PCR kit to be more than satisfactory.
7. The cycling profile is usually greatly modified, especially the primer extension step. As a general rule, allow 1 minute of primer extension at 68° for each 1000 bp to be amplified from the template.

Long-range technology can also be used to increase the sensitivity of standard PCR procedures (Fig. 18.16). When long-range reaction components are used to amplify typical length PCR products (200–900 bp), there is almost always an increase in sensitivity. When used in an RT-PCR format, the ability to detect very low abundance cDNAs improves. The day is coming when all routine PCR-based transcription assays will feature the two-enzyme format. There are several excellent kits available from biotech vendors, each of which features a unique enzyme blend and reaction chemistry that can support both long-range and short-range applications.

Single-cell PCR

Everyone wants to be able to amplify PCR products from a single cell, though most people would be happy if they could amplify something from a few cells. The good news is that the technology to amplify a single-copy gene

from a single cell is in place and in use in many places. This has profound implications in many facets of biology and medicine, especially in the areas of diagnostics and understanding normal development pathways and the progression of disease.

Some points to keep in mind: (1) well-characterized, PAGE- or HPLC-purified primers are very important for successful amplification of the intended product without competing spurious band formation; (2) individual cells can be isolated by limit dilution, microdissection, or by other methods that may be available; (3) standard PCR formulations can be used (primer concentration, dNTPs, Mg^{++}), though the addition of $1\text{ ng}/\mu\text{l}$ prokaryotic carrier RNA (e.g., *E.coli* tRNA or rRNA) is recommended to drive forward hybridization kinetics between primer and template; and (4) a hot-start approach is almost always required, and the inclusion of nested primers in a secondary reaction format as described above is very helpful in terms of amplification specificity.

Single-cell PCR has numerous advantages. First, the template material is completely unique to that cell. In contrast, nucleic acids purified from thousands or millions of cells are a heterogeneous mixture and at best represent an "average". Briefly, a single cell should be transferred to thin-wall PCR tube in a maximum of $1\times$ PCR reaction buffer, snap frozen on dry ice (do *not* place this tube in liquid nitrogen) and then stored at -70° for up to 6 months. When ready to perform PCR, slowly thaw the tube on ice and then immediately add sufficient reaction components to produce a $50\mu\text{l}$ reaction. Be sure to perform an initial melt of 5 minutes at 95° , followed by 40 amplification cycles. If no product is visible, transfer a $1\mu\text{l}$ aliquot into $49\mu\text{l}$ of fresh PCR reaction mix and perform an additional 40 cycles using either the same primers or a nested primer set. The only real disadvantage to single-cell PCR is that it is not practical when large numbers of cells need to be analyzed individually.

Analyzing transcriptional activity in a single cell (single-cell RT-PCR) is a bit more of a challenge, though it really isn't that difficult to do. Several interesting approaches have been described for the recovery of live cells (Tong *et al.*, 1994; Kelso *et al.*, 1999; Esumi *et al.*, 2006) as well as cells collected by laser capture from a tissue section (Zhan *et al.*, 2002). Some investigators perform a mini-RNA isolation procedure in a PCR tube containing a single-cell, while other investigators prefer to simply handle the single cell (which will have lysed upon thawing) as if it were a purified RNA prep: the sample is heat denatured in the presence of the downstream primer(s), after which RNasin, reverse transcriptase, and other reaction components are added. cDNA is synthesized for 30–60 minutes and, after heat inactivation of the reverse transcriptase, the cDNA is used as template for PCR either in a multiplex format or by splitting the cDNA into multiple tubes for analysis with different primer sets. It is not uncommon to perform 40–50 cycles. As suggested above, we have had excellent success detecting very low abundance transcripts by performing 35–40 cycles, transferring a $1\mu\text{l}$ aliquot into $49\mu\text{l}$ of fresh PCR master mix, and then performing additional cycles.

Splinkerette PCR

A splinkerette is a modified adapter molecule with a hairpin structure involving the 3' end of one of the two strands (Fig. 18.17a). One of the major applications of splinkerette PCR is chromosome walking in order to obtain upstream and downstream regions flanking a known sequence (Devon *et al.*, 1995). Better known techniques that are used for the same objective are inverse PCR and vectorette PCR (Riley *et al.*, 1990; Horn *et al.*, 2007).

The impetus behind the development of splinkerette PCR is a phenomenon known as end-repair priming. This occurs when the adapter- and insert DNA molecules that are not ligated undergo a filling-in reaction in which the 3' end(s) of molecules are extended in PCR cycle 1. Recessed 3' ends are present when restriction enzyme digestion produces fragments with a 5' overhang. These ends have a proclivity to anneal following template denaturation in PCR cycle 1, resulting in a competition reaction of sorts in which the amplification of non-specific, unwanted products detracts from the sensitivity of the assay. In the case of non-ligated splinkerette adapters, the 3' end of stand with the hairpin primes its own extension, resulting in a looped domain that is unable to support subsequent non-specific amplification (Fig. 18.17b). Thus, splinkerette PCR minimizes end-repair priming and, because the splinkerettes are not kinased, there is no opportunity for ligation between the lower stand of the splinkerette and the DNA template that is intended for amplification.

The splinkerette amplification strategy is attractive (1) when there is only enough sequence information to construct one gene-specific primer; (2) when there are issues with adapter self-ligation, especially when the DNA to be ligated to the adapter(s) has a 5' or 3' overhang; or (3) when non-specific priming is an issue. Moreover, splinkerette PCR does not require circularization of the molecules under investigation. For a review of this and similar strategies, see Hui *et al.* (1998).

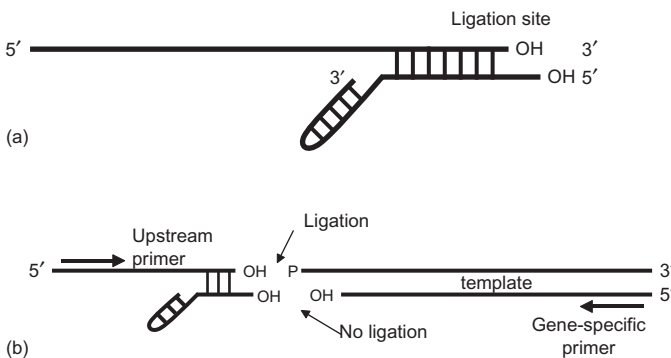


Figure 18.17 Splinkerette PCR. (a). Structure of a splinkerette adapter. Splinkerettes do not have a 5' phosphate, so only one of the two strands is able to be ligated to the intended template. (b). Formation of a non-amplifiable loop domain in non-ligated adapters minimizes the formation of spurious products due to 3' end-repair priming.

The hunt for alternative transcription start sites

An important contemporary question associated with gene regulation is the possibility that transcripts representing more than one transcription start site (TSS) may result from a single gene locus. This has already been observed in nature in both plant and animal species and appears to be an evolutionarily conserved process with an important role in gene expression. Most often, transcription from a particular genetic locus produces mRNAs representing a primary TSS and perhaps several nearby secondary TSSs corresponding to a single TATA box (Wray *et al.*, 2003). Reports of multiple TATA boxes within the same promoter, while rare, have been described in animal models (Fra *et al.*, 2000; Yamada *et al.*, 2001; reviewed by Farrell and Bassett, 2007).

In plant models, multiple TATA boxes initiating transcripts at different TSSs are less frequent than is observed in animals, though some interesting examples have been reported. For example, 5' RLM-RACE was used to demonstrate that a leucine-rich repeat receptor kinase gene (*inrpk1*) in morning glory (*Ipomoea [Pharbitis] nil*) produces transcripts from at least nine different TSSs using three different consensus and non-consensus TATA boxes (Bassett *et al.*, 2004), and does so in a tissue-specific manner. Similar types of gene regulatory events have been observed in *Arabidopsis* and tobacco (Dehesh *et al.*, 1994; Adams *et al.*, 1995). Thus, there are examples of multiple functional TATA elements just as there are examples of multiple promoters controlling certain nuclear genes, at least in animal models. The nuclear-encoded rat mitochondrial glycerol-3-phosphate dehydrogenase (G3PD) locus, for example, produces transcript variants encoding three different first exons due to the action of three distinct promoters (Weitzel *et al.*, 2000; Weitzel *et al.*, 2001).

This type of information is important because when the TSS changes, there may be an accompanying impact on the cognate protein, should translation occur. For example, the dehydrin 2 gene (*PpDhn2*) in peach (Wisniewski *et al.*, 2006) initiates from a consensus TATA box approximately 100 bp upstream of the translation start site in bark, while in mature fruit, initiation occurs an additional 125 bp upstream of the conventional ATG from an unidentified element which could be a nearby non-consensus TATA box. The resulting fruit-specific transcript contains an upstream AUG in good translational context, followed by an intron, which contains the consensus TATA box used in bark tissues. The result of splicing out of this intron places the upstream AUG in frame with the conventional translation start of the protein. This protein variant would therefore acquire an additional 34 amino acids at the N-terminus, and bioinformatics software has predicted this to be a mitochondrial transit sequence (Bassett *et al.*, 2009).

Classical RACE, as described above, has the unfortunate habit of producing many false positives as a consequence of the mechanics of the method itself, namely, that non-full-length cDNAs, often representing non-full-length mRNAs, are represented at the end of the experiment and must be characterized. 5' RLM-RACE, also described above, is now the definitive method for unambiguous TSS identification.

Protocol: first-strand cDNA synthesis

In advance: assess the quality of the RNA template by minigel electrophoresis.

1. Prepare master mix 1:

Component	Per reaction ¹²
H ₂ O	5 μ l
dNTP (10 mM)	1 μ l
Random primers (2.0 μ g/ μ l) ¹³	1 μ l
Oligo(dT) ₁₂₋₁₈ (1.0 μ g/ μ l)	1 μ l
Total	8 μ l

2. Distribute 8 μ l of master mix 1 into each sample tube.
3. Add 2 μ l of RNA (50–1000 ng) to each tube and do not vary the mRNA mass that is added to each tube.
4. Heat tubes to 70° for 5 min. At the conclusion of this RNA denaturation step, place samples on ice for at least 2 min.
5. While the RNA is being heat-denatured (step 4), prepare master mix 2:

Component	Each
H ₂ O	6 μ l
10 \times first-strand buffer ¹⁴	2 μ l
RNasin® (20 U/ μ l)	1 μ l
AMV R.T. (20 U/ μ l)	1 μ l
Total	10 μ l

6. Distribute 10 μ l of master mix 2 into each sample tube containing the denatured RNA.
7. Incubate at 25° for 5 min.

Note: This short room temperature incubation (1) facilitates primer annealing to the template, and (2) gives the reverse transcriptase an opportunity to extend the random primers and/or oligo(dT) for greater thermostability when the incubation temperature is increased in the following step. This almost always has the effect of producing much longer cDNA molecules.

¹²If preparing a master mix for multiple cDNA syntheses, be sure to multiply the “per reaction” volumes by one more than the total number of required reactions. This is a standard procedure to ensure an adequate volume of master mix in the event of subtle micropipetting errors. This same procedure also applies to the preparation of the PCR master mix.

¹³Alternatively, use a gene specific primer at a final concentration of about 1 μ M (required volume variable) instead of priming with oligo(dT) and/or random primers.

¹⁴Typical 10 \times first-strand buffer: 100 mM Tris, pH 8.3; 500 mM KCl; 40 mM MgCl₂. This buffer is usually provided with the reverse transcriptase.

8. If using a thermostable or enhanced reverse transcriptase, incubate samples at 50° for 50 min. Otherwise, incubate the sample at 37°–42° for 60 min.
- Note: There are numerous RT formulations that support first-strand cDNA synthesis at elevated temperatures. These are widely available from any of a number of vendors.*
9. Heat samples to 95° for 5 min to destroy the reverse transcriptase. Failure to do so can be detrimental to subsequent PCR-based procedures.
 10. The newly synthesized first strand cDNA is now ready for any one of several RT-PCR applications and should be placed on ice. Alternatively, cDNA can be stored at –20° until ready to use.

Protocol: PCR amplification of cDNA

1. Prepare PCR master. Typically, for each 50 μ l reaction to be performed, mix 5 μ l 10 \times PCR buffer (without Mg⁺⁺); 3 μ l 25 mM MgCl₂ (subject to change if optimization is required); 0.1 μ M upstream primer (volume variable); 0.1 μ M downstream primer (volume variable); 0.5 μ l Taq DNA polymerase (1.0–2.5U per reaction); water to 48 μ l.

Note 1: If different sequences are to be amplified in each reaction tube, then omit the primers from the master mix, adding each pair to the proper reaction tube directly. If the cDNA template is shared among all samples, it may be added directly to the PCR master mix.

Note 2: Never add all of cDNA to the master mix. By reserving a small aliquot, it will be possible to support several rounds of PCR from the same experimental source using various primers at a later date.

2. Aliquot 48 μ l of the master mix into each reaction tube.
3. Add 2 μ l of cDNA first-strand products to each reaction tube.
4. If using a thermal cycler without a heated lid, overlay 30 μ l of mineral oil onto each reaction. Omit mineral oil if using a thermal cycler with a heated lid.

Note: Keep in mind that the use of mineral oil interferes with the reproducibility of data.

5. Cycle the samples according to the following suggested generic profile. These parameters have been optimized for an Applied Biosystems 9700 thermal cycler using 500 μ l PCR tubes. Remember that cycling parameters are subject to optimization for each new set of primers, changes in the reaction composition, and model of thermal cycler.

Initial melt	94° for 2 min	
Cycling		
Denaturation	94° for 30 s	} 30 cycles
Primer annealing ¹⁵	55° for 30 s	
Primer extension	72° for 1 min	
Final extension	72° for 5 min	
Soak	4° until ready to use	

¹⁵The correct temperature at which to anneal (T_a) is a direct function of the T_m of the oligonucleotides being used. Consult the manufacturer's specification sheet that accompanied the primers. If uncertain where to begin, use $T_a = T_m - 5^\circ$. Recall that the T_m may be quickly estimated according to the formula $T_m = 2^\circ(A+T) + 4^\circ(G+C)$. Alternatively, the Applied Biosystems 9700 thermal cycler has a built-in utility for T_m calculation.

6. Remove an aliquot (5 μ l from a 50 μ l reaction) and analyze size and variety of products by minigel electrophoresis. PCR products are commonly analyzed in 2 to 3% agarose gels prepared in 1 \times TAE buffer, as described earlier in this chapter.

Cloning PCR products

The ability to amplify and reamplify DNA sequences by PCR has rendered routine vector-insert ligation something of a rarity. After all, if one requires more of the same product, it can be generated again by PCR using the same primers. Cloning of PCR products is commonly performed to (1) facilitate sequencing; (2) facilitate *in vitro* transcription; (3) physically separate a mixture of PCR products; and (4) support the expression of recombinant proteins. Of these, sequencing and *in vitro* transcription can be supported without cloning: reamplification of PCR products with modified up- and downstream primers consisting of appropriate 5' promoter sequences precludes the requirement for a vector. Further, multiple PCR products can be separated and recovered from agarose gels or directly from the reaction tube. Numerous kits are available PCR product clean-up, all of which are simple to use, expedient, and economical. For example, a gel slice containing a band of interest can be melted in a solution of guanidinium isothiocyanate (GTC) to create a chaotropic, high-salt environment. Alternatively, GTC can be added directly to the tube in which the cycling was performed. In either case, the PCR products will become bound to a silica matrix (Vogelstein and Gillespie, 1979). Residual reaction components, including primers, nucleotides, enzyme, and melted agarose, are washed away. The purified DNA is eluted from the silica matrix under very low-salt conditions, using 10 mM Tris, pH 8.0 or nuclease-free water.

Should cloning be deemed necessary, there are several options. *Taq* DNA polymerase exhibits a very useful habit of adding an extra nucleotide, most often an A, at the 3' end of each strand of the product that it generated. This is the basis for a convenient method known as TA cloning, in which the investigator anneals and then ligates the *Taq* A-tailed product into a vector that exhibits a 3' "T over-hang" on each strand (Fig. 18.18). These vectors are widely available and most have a least one selectable marker (e.g., Amp^r, Tet^r, or Kan^r) as well as a lacZ gene for blue/clear color selection.

There are several advantages associated with A-tailed products: (1) A-tailed products may be cloned directly without any knowledge of the base sequence of the PCR product; (2) A-tailed products eliminate the need for primers carrying 5'-overhanging restriction enzyme- or other sequences; (3) PCR products can usually be cloned directly, though purification can be helpful; (4) directional cloning is possible when hemiphosphorylated PCR products are generated (5' phosphorylation of one of the primers); and (5) vector cloning sites are usually located in a polylinker region (multiple cloning site, or MCS). This will facilitate excision using any of a number of restriction enzymes.

The ligation itself is mediated by T4 DNA ligase in a reaction requiring several hours. Note that some thermostable enzymes used for PCR demonstrate

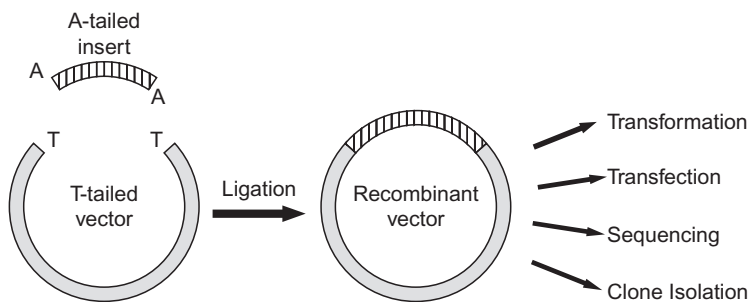


Figure 18.18 TA Cloning. PCR products generated with *Taq* polymerase exhibit a single base overhang, most often an adenine, at the 3' end of each strand. This facilitates ligation into a linearized vector that has been tailed with a thymine overhang (T-tailed). Such vectors are widely available. Cloning PCR products in general supports DNA sequencing, transformation into prokaryotes, transfection into eukaryotes for expression studies, and an overall strategy for physically separating multiple PCR products from each other.

a 3'→5' exonuclease (proofreading) activity and produce blunt-ended PCR products, which cannot be used for TA cloning unless they are tailed. This is easily accomplished by cleaning up a completed PCR reaction to remove the proofreading enzyme (PCR product cleanup kit), followed by incubation of the blunt-end PCR products with *Taq* polymerase (and other reaction components) for 15 minutes at 72°, followed by another clean-up procedure. Failure to perform this second clean-up procedure will reduce the efficiency of ligation in subsequent steps.

Protocol: a-tailing of blunt-end PCR products

1. Purify PCR products with a silica-based clean-up kit.
2. Prepare the tailing reaction:

Blunt-end PCR products	10.0 μl
H ₂ O	7.0 μl
10× Reaction Buffer (Mg ⁺⁺)	2.0 μl
10 mM dNTP mix	0.5 μl
<i>Taq</i> polymerase	0.5 μl
3. Incubate samples at 72° for 15 min.
4. Purify tailed PCR products as in step 1.
5. Store purified PCR products at -20°.

Protocol: TA cloning ligation reaction

Two or three ligation reactions for each PCR product should be assembled in order to maximize cloning efficiency: remember that the precise ratio of

vector to insert influences the ligation efficiency. Note that multiple ligations are suggested only for first-time users. It is generally necessary to perform only one ligation reaction because a large number of transformants (i.e., maximum ligation efficiency) is not required when only a single type of insert is expected.

1. Assemble the ligation reactions listed here. Be sure to add the enzyme last. As needed, the precise ratio of vector:insert can be modified to suit the investigator.

	10× Ligase buffer	T-tailed vector	PCR product	H ₂ O	T4 DNA ligase
Reaction 1	1 µl	10 ng	0.1 µl	<i>q.s.</i> 9 µl	1 µl
Reaction 2	1 µl	10 ng	1.0 µl	<i>q.s.</i> 9 µl	1 µl
Reaction 3	1 µl	10 ng	2.0 µl	<i>q.s.</i> 9 µl	1 µl

2. Mix the reaction(s) gently after addition of the ligase.
3. Incubate for a minimum of 3 h at 16°.

Note: Ligation reactions that are performed overnight generally produce the highest number of clones.

4. Thaw an aliquot of frozen (−80°) competent cells on ice.

Note: These cells may be purchased or prepared in advance. Be sure that the bacterial host strain is compatible with blue/clear color selection if using this selection scheme. Competent E. coli cells, strain JM109, works quite well in this application.

5. Pipette 20 µl aliquots of competent cells into prechilled microfuge tubes on ice.
6. Add 1 µl ligation reaction to each tube. Stir gently with pipette tip.
7. Incubate tubes on ice for 20 min.
8. Place the tubes in a 42° water bath for *exactly* 45 s. Do not shake these tubes.

Note: This heat shock is the most critical step of all, often determining the success or failure of the transformation. Do not heat at a higher temperature or for a longer period than recommended here.

9. Immediately return the tubes to the ice for 2 min.
10. Add 400 µl room temperature LB medium or room temperature SOC medium¹⁶.
11. Incubate at 37° for 1 h with shaking (200 rpm).
12. Plate an aliquot (10–100 µl) of each transformation reaction onto standard LB agar plates that have been supplemented with the appropriate antibiotic. Reserve the unused portion of the transformation by storing it at 4°.
13. Spread the transformation reaction aliquot over the surface of the agar as completely as possible using a sterile bent glass rod or similar implement.
14. Keep the plates upright for at least 5 min to allow all of the liquid to absorb into the agar.

¹⁶LB medium (per liter): 10 g tryptone, 5 g yeast extract, 10 g NaCl; autoclave. SOC medium (per liter): 20 g tryptone, 5 g yeast extract, 0.6 g NaCl, 0.19 g KCl, 2 g MgCl₂, 2.5 g MgSO₄, 3.6 g glucose; autoclave.

15. Invert the plates and incubate at 37° overnight.
 16. The following morning, remove the plates from the incubator as soon as possible. This will minimize the formation of satellite colonies.
 17. To test whether a colony on the dish is a true transformant, pick¹⁷ a small portion of the colony from the plate and inoculate the live bacteria directly into a 100 μ l PCR reaction mix: 10.0 μ l 10 \times PCR reaction buffer (containing Mg²⁺); 2 μ l 10 mM dNTP; 0.5 μ M downstream primer (volume variable); 0.5 μ M upstream primer (volume variable); 1 l Taq polymerase; water *q.s.* 98 μ l.
- Note 1: The heating steps associated with PCR will break open the bacterial cells and denature the DNA, thereby making the plasmid accessible to the primers.*
- Note 2: Do not transfer too much of the colony. Otherwise, amplification will be inhibited. Simply touching the surface of a colony with a toothpick is all that is needed.*
18. If necessary, overlay the reaction with mineral oil. Cycle with the same amplification parameters used to generate the PCR product.
 19. At the conclusion of the amplification, remove a representative aliquot of the reaction (10 μ l) and analyze by agarose gel electrophoresis. Colonies that contain a true insert will manifest a band identical in size to that observed for the PCR product prior to cloning.
 20. Return to any of the original colonies that generated a PCR product. Use the remainder of the colony to streak a new agar plate and to prepare a glycerol stock for long-term storage.

TOPO® Cloning

A novel, very widely respected alternative approach for cloning PCR products, or any double-stranded DNA, is the TOPO® Cloning method (Invitrogen). The key to TOPO Cloning is the use of the *Vaccinia* virus enzyme topoisomerase I. This enzyme functions both as a sequence-specific restriction enzyme (CCCTT) as well as a ligase. The method of cloning is rapid and efficient, often rendering >95% cloning efficiency in as few as 5 minutes and at room temperature. The TOPO Cloning method is compatible with TA-cloning, blunt-end ligations, and directional cloning.

There are several TOPO vectors that support a variety of cloning formats. In each case the TOPO vector has already been linearized and the topoisomerase I enzyme is already attached via a 3' phosphotyrosyl bond (Shuman, 1994); such vectors are said to be "activated" and will not exhibit self-ligation (Fig 18.19). Successful ligation to a TOPO vector requires that the intended DNA insert exhibit free 5' -OH ends, meaning that PCR products must be generated with primers that do not have 5' phosphate groups. If the cloning of cDNA, genomic DNA, or restriction fragments is intended, the DNA must be dephosphorylated first, which is itself a rather simple manipulation. The typical ligation incubation

¹⁷There have been numerous reports of PCR inhibitors found in some brands of wooden toothpicks. It would be judicious to test toothpicks for reaction inhibition before using them to pick experimental clones. A better, simple solution is to use plastic toothpicks or one end of a sterile, disposable inoculating loop.

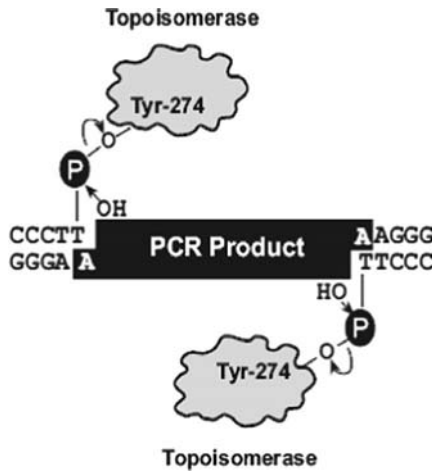


Figure 18.19 TOPO[®] Cloning. TOPO vectors are linearized and already have topoisomerase I attached. This enzyme is linked to the vector via a 3' phosphate. The PCR product, i.e., the insert, contributes at 5' -OH. Courtesy of Invitrogen.

period using a TOPO vector is remarkably short, on the order of 5 min. Upon ligation, the topoisomerase I enzyme is released from the now recombinant plasmid construction. Bacterial transformation proceeds as usual.

Other amplification procedures

PCR is by far the most widely known and used nucleic amplification procedure. It is not, however, the only amplification method. PCR has its shortcomings, and some of the alternative amplification procedures circumvent these shortcomings and, in some cases, rival or exceed the amplification potential of PCR.

Linear RNA amplification (Eberwine process)

Microarray analysis is a procedure that suffers when PCR-amplified sequences are hybridized to the array. This is a consequence of the bias that PCR has toward certain molecules, resulting in a distortion of the true abundance relationships among the various molecular species after even a few cycles. The need for any amplification at all occurs when only very small quantities of RNA are available (typically nanogram quantities) from a biological source, because microarray experiments require at least 5–10 μg RNA.

The preferred method for generating more RNA from a paucity of template material is known as the Eberwine process or, more commonly, as linear RNA amplification (Van Gelder *et al.*, 1990). As shown in Fig. 18.20, the process involves converting RNA into double-stranded cDNA, after which the cDNA

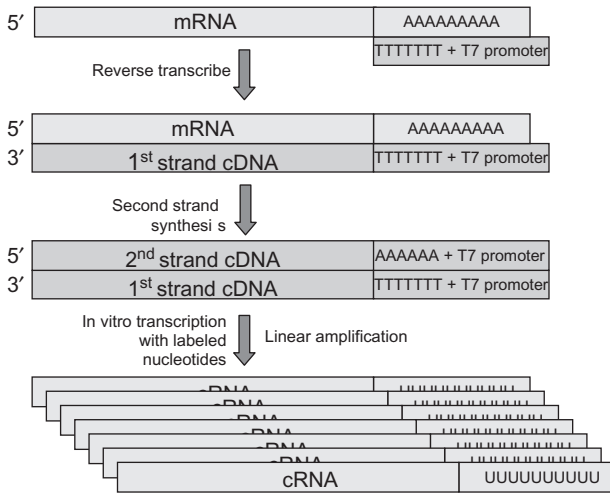


Figure 18.20 Eberwine linear amplification process. In order to produce more RNA while preserving the abundance relationships among transcript species, cDNA is synthesized through using a downstream primer-adaptor that encodes an *in vitro* transcription promoter. Double-stranded cDNA is then transcribed, often in the presence of label, to produce cRNA which is then used for array hybridization.

is transcribed to produce more RNA. Because the cDNA directly reflects the abundance of the RNA, subsequent transcription of the cDNA maintains the relative abundance of the resulting RNA while providing more of it. The incorporation of a transcription promoter via a downstream primer-adaptor, supports the synthesis of labeled complementary RNA (cRNA), which can then be hybridized to an array. This procedure may be used to generate sense (+) or antisense (–) transcripts, depending of the requirements of the investigator and the orientation of the promoter(s). The overall bias using this method of amplification is quite low, by comparison with PCR.

Strand displacement amplification

SDA (strand displacement amplification) is a patented isothermal, *in vitro* method for sequence amplification which was developed by Becton Dickinson (Walker *et al.*, 1992a; 1992b). The method routinely produces at least 10^8 -fold amplification, and much higher levels of amplification have been reported. SDA typically begins with a target generation step in which template molecules that are flanked by a restriction enzyme site are produced. The exponential amplification stage follows. Succinctly, the method is based on a restriction endonuclease cutting its recognition site which then allows a polymerase to extend the nick from the 3' end. This has the effect of displacing the downstream strand. The displaced (+) strand serves as target for (–) strand reaction. Repeated nicking, strand displacement, and annealing of primers to the displaced strands results in exponential growth.

Nucleic acid sequence based amplification (NASBA)

NASBA (Compton, 1991) is a primer-dependent isothermic process for RNA synthesis process, the patent for which is currently owned by bioMérieux, Inc. It is also commonly referred to as “self-sustained sequence replication” whereby RNA is converted into cDNA, which is transcribed to produce more RNA, all at 41°. NASBA routinely produces a high level of amplification, on the order to 10⁹. Reports from many sectors suggest that this amplification method is at least as sensitive as PCR, if not more so. NASBA is a widely respected molecular diagnostics technique, particularly for retroviral detection.

Ligase chain reaction

The ligase chain reaction (LCR) is an amplification process that differs from PCR in that it involves a thermostable ligase to join two probes or other molecules together which can then be amplified by standard PCR cycling (Barany, 1991). Thus, LCR requires two completely different enzymes to operate properly: ligase, to join probe molecules together, and a thermostable polymerase (e.g., *Taq* polymerase) to amplify those molecules involved in successful ligation. The probes involved in the ligation are designed such that the 5' end of one probe is directly adjacent to the 3' end of the other probe, thereby providing the requisite 3'-OH and 5'-PO₄ group substrates for the ligase.

LCR was originally developed to detect point mutations; a single base mismatch at the junction of the two probe molecules is all that is need to prevent ligation. By performing the ligation right at the T_m of the oligonucleotide probe, only perfectly matched primer:template duplexes will be tolerated. LCR can also be used to amplify template molecules that have been successfully ligated for the purpose of assessing ligation efficiency and producing a large amount of product with even greater specificity than PCR. Thus, LCR is not necessarily an alternative, but rather a complement, to PCR.

References

- Aatsinki, J.T., J.T. Lakkakorpi, E.M. Pietila and H.J. Rajaniemi. (1994). A coupled one-step reverse transcription PCR procedure for generation of full-length open reading frames. *BioTechniques* **16**, 282–285.
- Adams, E., L. Kozma-Bognar, G. Dallmann and F. Nagy. (1995). Transcription of tobacco phytochrome-A genes initiates at multiple start sites and requires multiple cis-acting regulatory elements. *Plant Mol. Biol.* **29**, 983–993.
- Alexander, M., M. Denaro, R. Galli, B. Kahn and N. Nasrin. (1990). Tissue-specific gene regulation of the glyceraldehyde-3-phosphate-dehydrogenase gene by insulin correlates with the induction of an insulin-sensitive transcription factor during differentiation of 3T3 adipocytes. In *Obesity: Toward a Molecular Approach* (G. Bray, D. Ricquier and B. Speigleman, Eds). A.R. Liss, New York, NY. pp. 247–261.
- Balajee, A.S., A. May and V.A. Bohr. (1999). DNA repairs of Pyrimidine dimers and 6-4 photoproducts in the ribosomal DNA. *Nucleic Acids Res.* **27**(12), 2511–2520.

- Barany, F. (1991). Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci* **88**, 189–193.
- Bassett, C.L., M.E. Wisniewski, T.S. Artlip, G.A. Richart, J.L. Norelli, and R.E. Farrell, Jr. (2009). Comparative expression and transcript initiation of three peach dehydrin genes. *Planta* **230**, 107–118.
- Bassett, C.L., M.L. Nickerson, R.E. Farrell, Jr. and M. Harrison. (2004). Multiple transcripts of a leucine-rich repeat receptor kinase from morning glory originate from different TATA boxes in a tissue-specific manner. *Mol. Gen. Genom.* **271**(6), 752–760.
- Bassett, C.L., M.L. Nickerson, R.E. Farrell, Jr., T.S. Artlip, A.E. Ghaouth, C.L. Wilson and M.E. Wisniewski. (2005). Characterization of an S-locus receptor protein kinase-like gene from peach. *Tree Physiol.* **25**, 403–411.
- Bhatia, P., W.R. Taylor, A.H. Greenberg and J.A. Wright. (1994). Comparison of glyceraldehyde-3-phosphate-dehydrogenase and 28S ribosomal RNA gene expression as RNA loading controls for Northern blot analysis of cell lines of varying malignant potential. *Analyt. Biochem.* **216**, 223–226.
- Bonini, J.A. and C. Hofmann. (1991). A rapid, accurate nonradioactive method for quantitating RNA on agarose gels. *BioTechniques* **11**, 708–710.
- Chretien, S., A. Dubart, D. Beaupain, N. Raich, B. Grandchamp, J. Rosa, M. Goossens and P.H. Romeo. (1988). Alternative transcription and splicing of the human porphobilinogen deaminase gene results either in tissue-specific or in housekeeping expression. *Proc. Natl. Acad. Sci*, **85**, 6–10.
- Cheng, S., C. Fockler, W.M. Barnes and R. Higuchi. (1994). Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci* **91**, 5695–5699.
- Chou, Q. (1992). Minimizing deletion mutagenesis artifact during *Taq* DNA polymerase PCR by *E. coli* SSB. *Nucleic Acids Res.* **20**, 4371.
- Compton, J. (1991). Nucleic acid sequence-based amplification. *Nature* **350**, 91–92.
- Danielson, P., S. Forss-Petter, M. Brow, L. Calavette, J. Douglas, R. Milner and J. Sutcliffe. (1988). A cDNA clone of the rate mRNA encoding cyclophilin. *DNA* **7**, 261–267.
- Dehesh, K., C. Franci, R.A. Sharrock, D.E. Somers, J.A. Welsch and P.H. Quail. (1994). The *Arabidopsis* phytochrome A gene has multiple transcription start sites and a promoter sequence motif homologous to the repressor element of monocot phytochrome A genes. *Photochem. Photobiol.* **59**, 379–384.
- de Leeuw, W., P. Slagboom and J. Vijg. (1989). Quantitative comparison of mRNA levels in mammalian tissues: 28S ribosomal RNA level as an accurate internal control. *Nucleic Acids Res.* **17**, 10137–10138.
- Devon, R.S., D.J. Porteous and A.J. Brookes. (1995). Splinkerettes – improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res.* **23**, 1644–1645.
- Dieffenbach, C. and G.S. Dveksler (Eds). (2003). *PCR Primer*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Dodge, G.R., I. Kovalszky, J.R. Hassell and R.V. Iozzo. (1990). Transforming growth factor β alters the expression of heparin sulfate proteoglycan in human colon carcinoma cells. *J. Biol. Chem.* **265**, 18023–18029.
- Esumi, S., R. Kaneko, Y. Kawamura and T. Yagi. (2006). Split single-cell RT-PCR analysis of Purkinje cells. *Nature Protocols* **1**, 2143–2151.
- Fahle, G.A., V.J. Gill and S.H. Fischer. (1999). Optimal activation of isoposalen to prevent amplicon carryover. *J. Clin. Microbiol.* **37**, 261–262.

- Farrell, R.E. and C.L. Bassett. (2007). Multiple transcript initiation as a mechanism for regulating gene expression. In *Regulation of Gene Expression in Plants* (C.L. Bassett Ed.). Springer Science, New York, NY. pp. 39–66.
- Farrell, R.E. and J.J. Greene. (1993). Regulation of *c-myc* and *c-Ha-ras* oncogene expression by cell shape. *J. Cell. Physiol.* **153**, 429–435.
- Feroze-Merzoug, F., I.M. Berquin, J. Dey and Y.Q. Chen. (2002). Peptidylprolyl isomerase A (PPIA) as a preferred internal control over GAPDH and β -actin in quantitative RNA analyses. *BioTechniques* **32**, 776–782.
- Fra, A.M., E. Pasqualetto, M. Mancini and R. Sitia. (2000). Genomic organization and transcriptional analysis of the human genes encoding caveolin-1 and caveolin-2. *Gene* **243**, 75–83.
- Frohman, M.A., M.K. Dush and G.R. Martin. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci.* **85**, 8998–9002.
- Frohman, M.A. (1995). Rapid Amplification of cDNA Ends. In *PCR Primer* (C.W. Dieffenbach and G.S. Dveksler, Eds). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. pp. 381–409.
- Gubler, U. and B.J. Hoffman. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.
- Gyllensten, U.B. and H.A. Erlich. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci.* **85**, 7652–7656.
- Haendler, B. and E. Hofer. (1995). Characterization of the human cyclophilin gene and of related pseudogenes. *Eur. J. Biochem.* **190**, 477–482.
- Henke, W., K. Herdel, K. Jung, D. Schnorr and S.A. Loening. (1997). Betaine improves PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* **25**, 3957–3958.
- Horikoshi, T., K.D. Danenberg, T.H. Stadlbauer, M. Volkenandt, L.C. Shea, K. Aigner, B. Gustavsson, L. Leichman, R. Frosing and M. Ray. (1992). Quantitation of thymidylate synthase, dihydrofolate reductase, and DT-diaphorase gene expression in human tumors using the polymerase chain reaction. *Cancer Res.* **52**, 108–116.
- Hui, E., P. Wang and S.J. Lo. (1998). Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell. Mol. Life Sci.* **54**, 1403–1411.
- Innis, M.A., D.H. Gelfand and J.J. Sninsky. (1999). *PCR Applications: Protocols for Functional Genomics*. Academic Press, San Diego, CA.
- Isaacs, S.T., J.W. Tessman, K.C. Metchette, J.E. Hearst and G.D. Cimino. (1991). Post-PCR sterilization: development and application to an HIV-1 diagnostic assay. *Nucleic Acids Res.* **19**, 109–116.
- Kawasaki, E.S. (1990). Amplification of RNA. In *PCR Protocols: A Guide to Methods and Applications* (M.A. Innis, D.H. Gelfand, J.J. Sninsky and T.J. White, Eds). Academic Press, San Diego, CA. pp. 21–27.
- Kelso, A., P. Groves, L. Ramm and A.G. Doyle. (1999). Single-cell analysis by RT-PCR reveals differential expression of multiple type 1 and 2 cytokine genes among cells within polarized CD4+ T cell populations. *Internat. Immunol.* **11**, 617–621.
- Kinoshita, T., J. Imamura, H. Nagai and K. Shimotohno. (1992). Quantification of gene expression over a wide range by the polymerase chain reaction. *Analyt. Biochem.* **206**, 231–235.
- Lau, W.Y., P.B. Lai, M.F. Leung, N. Wong, G. Chen, T.W. Leung and C.T. Liew. (2000). Differential gene expression of hepatocellular carcinoma using cDNA microarray analyses. *Oncol. Res.* **12**, 59–69.

- Leof, E.B., J.A. Proper, M.J. Getz and H.I. Moses. (1986). Transforming growth factor type beta regulation of actin mRNA. *J. Cell Physiol.* **127**, 83–88.
- Linton, J.P., J.J. Yen, E. Selby, Z. Chen, J.M. Chinsky, K. Liu, R.E. Kellems and G. F.Crouse. (1989). Dual bidirectional promoters at the mouse *dhfr* locus: cloning and characterization of two mRNA classes of the divergently transcribed Rep-1 gene. *Mol. Cell. Biol.* **9**(7), 3058–3072.
- Loh, E.L., J.F. Elliott, S. Cwirla, L.L. Lanier and M.M. Davis. (1989). Polymerase chain reaction with single sided specificity: analysis of T cell receptor delta chain. *Science* **243**, 217–220.
- Mansur, N., K. Meyer-Siegler, J. Wurzer and M. Sirover. (1993). Cell cycle regulation of the glyceraldehyde-3-phosphate-dehydrogenase/uracil DNA glycosylase gene in normal human cells. *Nucleic Acids Res.* **21**, 993–998.
- Maruyama, K. and S. Sugano. (1994). Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**, 171–174.
- McPherson, M.J. and B.D. Hames (Eds). (1995). *PCR 2: A Practical Approach*. IRL Press at Oxford University Press, Oxford.
- Mullis, K.B. and F.A. Faloona. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335–350.
- Mullis, K.B., F.A. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 263–273.
- Ochman, H., A.S. Gerber and D.L. Hartl. (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–625.
- Oleary, K.A., T.W. Beck and C.B. Kasper. (1994). NADPH cytochrome P-450 oxidoreductase gene: identification and characterization of the promoter region. *Biochem. Biophys.* **310**, 452–459.
- Olive, D.M., M. Simsek and S. Al-Mufti. (1989). Polymerase chain reaction assay for detection of human cytomegalovirus. *J. Clin. Microbiol.* **27**, 1238–1242.
- Paabo, S., J.A. Gifford and A.C. Wilson. (1988). Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Res.* **16**, 9775–9787.
- Pernas-Alonso, R., F. Morelli, U. di Porzio and C. Perrone-Capano. (1999). Multiplex semi-quantitative reverse transcriptase-polymerase chain reaction of low abundance neuronal mRNAs. *Brain Res. Protocols* **4**, 395–406.
- Rees, W.A., T.D. Yager, J. Korte and P.H. Von Hippel. (1993). Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* **32**, 137–144.
- Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner and S. Powell. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18**, 2887–2890.
- Robert, C., S. McGraw, L. Massicotte, M. Pravetoni, F. Gandolfi and M.-A. Sirard. (2002). Quantification of housekeeping transcript levels during the development of bovine preimplantation embryos. *Biol. Reprod.* **67**, 1465–1472.
- Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich and N. Arnheim. (1985). Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354.
- Sarkar, G., S. Kapelner and S.S. Sommer. (1990). Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Res.* **18**, 7465.
- Sellner, L.N., R.J. Coelen and J.S. Mackenzie. (1992). Reverse transcriptase inhibits Taq polymerase activity. *Nucleic Acids Res.* **20**, 1487–1490.

- Sellner, L.N. and G.R. Turbett. (1998). Comparison of three RT-PCR methods. *BioTechniques* 25(2), 230–234.
- Shaefer, B.. (1995). Revolution in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Analyt. Biochem* 227, 255–273.
- Shuman, S.. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J. Biol. Chem.* 269, 32678–32684.
- Solomon, D.H., K. O'Driscoll, G. Sosne, I.B. Weinstein and Y.E. Cayre. (1991). 1 alpha,25-dihydroxyvitamin D3-induced regulation of protein kinase C gene expression during HL-60 cell differentiation. *Cell Growth Different.* 2, 187–194.
- Spanakis, E.. (1993). Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res.* 16, 3809–3819.
- Steen, A.M., H. Luthman, D. Hellgren and B. Lambert. (1990). Levels of hypoxanthine phosphoribosyltransferase RNA in human cells. *Exp. Cell Res.* 186, 236–244.
- Tepper, C.G., M.M. Pater, A. Pater, H. Xu and G.P. Studzinski. (1992). Mitochondrial nucleic acids as internal standards for blot hybridization analyses. *Analyt. Biochem.* 203, 127–133.
- Tong, J., S. Bendahhou, H. Chen and W.S. Agnew. (1994). A simplified method for single-cell RT-PCR than can detect and distinguish genomic DNA and mRNA transcripts. *Nucleic Acids Res.* 22, 3253–3254.
- Triglia, T., M.G. Peterson and D.J. Kemp. (1988). A procedure to *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.* 16, 8186.
- Van Gelder, R.N., M.E. von Zastrow, A. Yool, W.C. Dement, J.D. Barchas and J.H. Eberwine. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci.* 87, 1663–1667.
- Veres, G., R.A. Gibbs, S.E. Scherer and C.T. Caskey. (1987). The molecular basis of the sparse fur mouse mutation. *Science* 237, 415–417.
- Vogelstein, B. and D. Gillespie. (1979). Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci* 76(2), 615–619.
- Voges, D., P. Zwickl and W. Baumeister. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Ann. Rev. Biochem.* 68, 1015–1068.
- Walker, G.T., M.C. Little, J.G. Nadeau and D.D. Shank. (1992a). Isothermal *in vitro* amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc. Natl. Acad. Sci.* 89, 392–396.
- Walker, G.T., M.S. Fraiser, J.L. Schram, M.C. Little, J.G. Nadeau and D.P. Malinowski. (1992b). Strand displacement amplification – an isothermal, *in vitro* DNA amplification technique. *Nucleic Acids Res.* 20, 1691–1696.
- Weitzel, J.M., S. Grott, C. Radtke, S. Kutz and H.J. Steitz. (2000). Multiple promoters direct the tissue-specific expression of rat mitochondrial glycerol-3-phosphate dehydrogenase. *Biol. Chem.* 381, 611–614.
- Weitzel, J.M., S. Kutz, C. Radtke, S. Grott and H.J. Steitz. (2001). Hormonal regulation of multiple promoters of the rat mitochondrial glycerol-3-phosphate dehydrogenase gene. *Eur. J. Biochem.* 268, 4095–4103.
- Winship, P.R. (1989). An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.* 17, 1266.
- Wisniewski, M.E., C.L. Bassett, J. Renaut, R.E. Farrell, Jr., T. Tworkoski and T.S. Artlip. (2006). Differential regulation of two dehydrin genes from peach (*Prunus*

- persica L. Batsch) by photoperiod, low temperature, and water deficit. *Tree Physiol.* **26**, 575–584.
- Wray, G.A., M.W. Hahn, E. Abouheif, J.P. Balhoff, M. Pizer, M.V. Rockman and L.A. Romano. (2003). The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**, 1377–1419.
- Yamada, M., A. Ozawa, S. Ishii, N. Shibusawa, T. Hashida, T. Ishizuka, T. Hosoya, T. Monden, T. Satoh and M. Mori. (2001). Isolation and characterization of the rat Prolactin-releasing peptide gene: multiple TATA boxes in the promoter region. *Biochem. Biophys. Res. Comm.* **281**, 53–56.
- Zhan, G., F. Shaheen, M. Mackiewicz, P. Fenik and S.C. Veasy. (2002). Single cell laser dissection with molecular beacon polymerase chain reaction identifies 2A as the predominant serotonin receptor subtype in hypoglossal motoneurons. *Neuroscience* **113**, 145–154.

19 Quantitative PCR Techniques

Rationale	449	Competitive PCR: major steps involved	477
Sensitivity index	450	Alternative approach: non-real-time competitive PCR	479
Quantitative approaches	451	Protocol: competitive PCR	480
Real-time PCR	453	Synthesis of non-homologous competitor	480
Real-time PCR platforms	456	Synthesis of first-strand cDNA	481
SYBR green assay	457	Competitive PCR (primary amplification)	482
TaqMan® assay	458	Competitive PCR (secondary amplification)	484
Molecular beacons	458	Image analysis considerations	485
LUX™ primers	460	Troubleshooting quantitative PCR techniques	486
Scorpions	461	References	488
Melting curve analysis	462		
Internal controls	463		
Exogenous controls	463		
Control reaction formats	467		
Negative control considerations	471		
Competitive PCR: key considerations	472		

Rationale

The polymerase chain reaction (PCR) is a widespread molecular technique that has revolutionized every aspect of biotechnology since its inception. In short, it is a primer-mediated, enzymatic amplification of specific genomic DNA or cDNA sequences. The successive application of the very familiar format of sample denaturation, primer annealing, and primer extension over 25–30 cycles results in the exponential amplification of target molecules. Readers who are not yet familiar with the mechanics of RNA-based PCR are strongly encouraged to read Chapter 18 in its entirety before moving forward.

PCR offers unsurpassed amplification potential. For simple “yes or no”-type experiments there are less stringent requirements for the input mass and purity of RNA compared to more traditional methods. This is true partly because PCR primers “frame” the domain that will be amplified, making PCR a bit more tolerant of lesser quality RNA than traditional cDNA synthesis and cloning procedures. For quantification, however, input mass of RNA, transcript quality, and overall amplification efficiency are key parameters.

Historically, the major difficulties associated with the synthesis of cDNA, both for cloning purposes and as an indirect measure of transcript abundance, have been (1) the inability to generate near full-length, *representative* cDNA molecules; and (2) the overall low efficiency of the process of reverse transcription itself. Currently, the extent to which these difficulties can be overcome is related directly to (1) the quality of the RNA template; (2) the quantity of the RNA template; and (3) the cDNA synthesis methodology, which is discussed in Chapter 17.

These days, the achievable level of sensitivity in PCR-based transcription assays is extraordinary. Each month, innovations appear in the literature that suggest new or improved approaches for maximizing sensitivity. As close as we might come, however, there is no such thing as absolute quantification. Human error, no matter how small, will always be a confounding variable in the unending quest for absolute accuracy.

Sensitivity index

This laboratory guide presents a number of methodologies for the isolation and subsequent characterization of RNA. It should be abundantly clear that each of these methods offers distinct advantages and disadvantages which extend into the realm of the overall reliability and sensitivity of the particular assay. Failure to adhere precisely to the prescribed methods or to include appropriate controls will absolutely have a negative impact on the sensitivity of an experiment and the interpretation of the resulting data.

Most often, levels of RNA in an experimentally related set of samples are described in terms of relative abundance, meaning that one of the samples is designated as the control or reference sample and to which all other samples are compared. Typically, an observed hybridization signal or the intensity of a PCR product is assigned a value of 1.0 and the behavior of each of the other samples is compared to this reference or baseline value. Be aware, however that relative abundance is directly related to the sensitivity of the assay being performed; evaluating the same set of samples using a battery of different methods or even the same method on different days will often produce startlingly different results.

The relative sensitivity of several common transcription assay methods is herein described in terms of what this Author affectionately calls "Farrell's RNA sensitivity index." As shown below, RNA-based assays for transcript quantification are arranged vertically, moving down from highest sensitivity to lowest. The sensitivity index ranking is based on three factors: intrinsic difficulty of the technique, reproducibility, and current average level of sensitivity.

It should be clear from reading earlier parts of this text that techniques which rank lower on the sensitivity index are nonetheless able to provide useful information. For example, the time-honored Northern analysis provides information about the native size of the mRNA produced at a specific locus, a datum not associated with methods of higher sensitivity ranking.

Highest Sensitivity
Real-Time PCR
End-Point PCR
Microarray Technology
Nuclear Runoff Assay
RNase Protection Assay
S1 Nuclease Protection Assay
Northern Analysis
Dot Blot Analysis
Lowest Sensitivity

Quantitative approaches

Quantitative approaches of all persuasions have been developed for the purposes of quantifying transcript abundance and determining gene copy number: the former is the focus of this chapter. Historically, quantification has been something of a challenge because of sample-to-sample variations in the efficiency of cDNA synthesis often due to residual chemistries such as ethanol, isopropanol, or phenol used to isolate the RNA from its biological source or, very commonly, due to excess salt. A primary concern among investigators wishing to measure transcript abundance by RT-PCR is the realization that the final mass of the PCR product(s) is dependent on the efficiency of the reverse transcriptase *and* the efficiency of the thermostable enzyme(s) used to support PCR, even if said actions occur in the same tube and by the action of the same enzyme. Thus, any reliable quantification method must factor in the relative efficiencies of both enzyme-mediated reactions because the overall efficiency of a technique is the product of the efficiencies of the individual steps involved. Keep in mind that the “efficiency index” is also user-dependent. At present, real-time PCR, in all of its various manifestations, is widely recognized as the sensitivity gold standard. Alternative approaches are also discussed below.

In order to ensure meaningful comparison between all samples, the process of normalization begins very early in an investigation, and there are a number of ways to approach this challenge. With respect to the reverse transcriptase reaction products, normalization of cDNA samples in days past has been based on sample input mass, performed most often (1) by preparing uniform dilutions of a single aliquot of the cDNA; (2) based on A_{260} measurements; or (3) based on amount (cpm) of incorporated radiolabel in instances where cDNA was performed in the presence of a small amount of isotopic tracer. Although these methods offer a reasonably accurate glimpse of the cellular biochemistry, a more reliable normalization procedure is to measure some type

of control or reference sequence, comparing its behavior to that of the experimental sequences under investigation. The inclusion of such a control is of prime importance because of the variegated nature of cDNA synthesis, with efficiencies commonly ranging from 10 to 90%. As discussed below, the reference transcript can be an endogenous mRNA or an artificial transcript of precisely known mass that is spiked into the biological sample prior to first-strand cDNA-synthesis. Under ideal conditions, PCR amplification should influence both control and experimental templates equally, when the reactions are properly formatted, which is the rationale for closely monitoring a control or reference transcript.

With respect to RT-PCR, inasmuch as PCR is an exponential process, the degree to which template is amplified is described by the equation

$$N = N_0 (1 + \text{eff})^n$$

where

N is the number of copies of PCR product at the end of amplification

N_0 is the initial number of copies of template

eff is the efficiency of the reaction, expressed as a decimal

n is the number of cycles performed.

It is important to note that this mathematical model can be used only to make an estimate of the amplification potential of PCR in a particular reaction tube because the efficiency of *each* round of “denature–anneal–extend” changes with each passing cycle, particularly in the latter cycles when the amplification efficiency of each cycle will plummet. In great measure, this is due to what is known as the plateau effect (Fig. 19.1). PCR product accumulates exponentially in the early stages of a reaction and the high-precision nature of the reaction during that time makes the data very reliable. As PCR

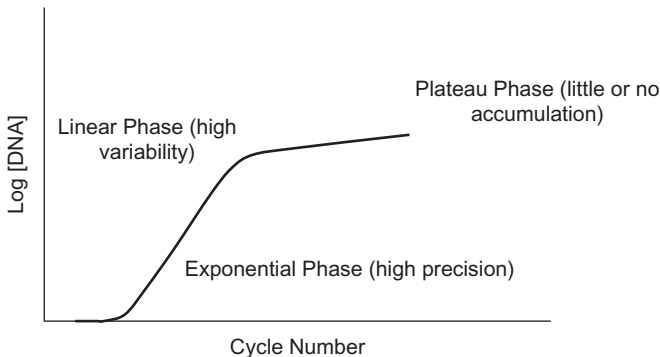


Figure 19.1 PCR plateau effect. The greatest precision in quantification results when product accumulation measurements and comparisons are made very early in the exponential phase. As the reaction progresses, the true abundance relationships among samples are lost because the reaction is becoming less efficient with each passing cycle.

progresses, the enzyme(s) is becoming less efficient (they are not thermostable forever), the nucleotide concentration is diminishing, as is the primer concentration, and there is an increasing number of new template molecules after each passing cycle which are competing for the reduced amounts of the primers, nucleotides, and polymerase. Thus, the reaction enters the linear phase in which there is great variability with respect to reaction efficiency. Finally, in the plateau phase, there is little, if any, product accumulation with each passing cycle, meaning that additional cycles within the plateau phase are pointless.

Why not, therefore, just run PCR for 8–12 cycles and then run an aliquot on a gel? The problem is that there is almost never enough product after only a dozen or so cycles to see anything on a gel. Plus, the opening the tube provides yet another opportunity for amplicon contamination. An older attempt to circumvent the product mass issue was to add ^{32}P to the reaction; the resulting PCR products could then be run on a gel and quantified by autoradiography. The problems with this method are two-fold: first, there is the potential for a huge radioactive mess, and second, the X-ray film has a narrow linear response, meaning the film has a limited ability to record the true abundance differences among samples. For these and other reasons, real-time PCR has been widely embraced.

Real-time PCR

A major breakthrough in gene- and transcript quantification was the development of what is commonly known as “real-time” PCR (Higuchi *et al.*, 1992; 1993). In this fluorescence-based approach, PCR products are quantified as they accumulate, as opposed to performing a fixed number of cycles and then assessing the amount and variety of products by electrophoresis, now commonly referred to as end-point analysis. Of late, many reports in the scientific literature refer to real-time PCR as “the new RT-PCR”. This nomenclature often leads to confusion, however, because of the immediate association of RT-PCR with the use of reverse transcriptase to synthesize cDNA, and the amplification thereof, which may not involve real-time PCR at all. Real-time PCR should be referred to as “real-time PCR” so that there is no question in the mind of the reader as to precisely which molecular techniques have been performed.

Real-time PCR requires some type of fluorescent labeling reagent in the reaction, a thermal cycler equipped with a fluorescence excitation device to induce the fluorescence and a sensitive fiberoptics system for fluorescence detection, all of which are controlled by a computer (Fig. 19.2). The several real-time PCR variants may involve the reaction tubes, the reaction format, the amplification and detection chemistries, or the design and functionality of the thermal cycler instrumentation. Succinctly, transcript abundance is extrapolated from the increasing level of fluorescence that is generated during PCR; these data are collected once in each PCR cycle and then related to a standard curve for extremely accurate quantification. The fluorescence may be the result of



Figure 19.2 LightCycler® 2.0 real-time PCR system shown with the top open. The instrument holds up to 32 samples in specially designed glass capillaries that travel on a rotating carousel during the PCR process. Courtesy of Roche Applied Science.

chemically modified oligonucleotides acting as primers or as probes, or by direct binding of a fluorescent dye to DNA molecules as they are synthesized.

The real-time approach proffers several advantages over traditional methods for measuring transcript abundance including greater speed because this is largely an automated process, and a wide dynamic range which often spans eight to nine orders of magnitude. The ability to recognize subtle variations in gene expression is due to the fact the determination of the original template mass is made early in the exponential phase of cDNA amplification. With respect to concerns regarding the amplified DNA, the closed-tube system reduces the risk of laboratory-wide amplicon contamination and data variations due to sample handling, since there is no post-PCR handling of samples. The accompanying software is capable of making sophisticated measurements; for example, melting curve analysis routinely performed which allows discrimination between true PCR products and primer-dimer artifacts.

The basic unit of measure in real-time PCR is the threshold cycle (C_T). This is the cycle at which the fluorescence emanating from a sample, due to PCR product accumulation, crosses a threshold level (e.g., fluorescence detected above background) that is pre-determined by the user. The higher the amount of template at the onset of a reaction, the sooner a significant fluorescence increase will be detected (Fig. 19.3). The important points to keep in mind are (1) low C_T = high template abundance; and (2) early detection is best, so that

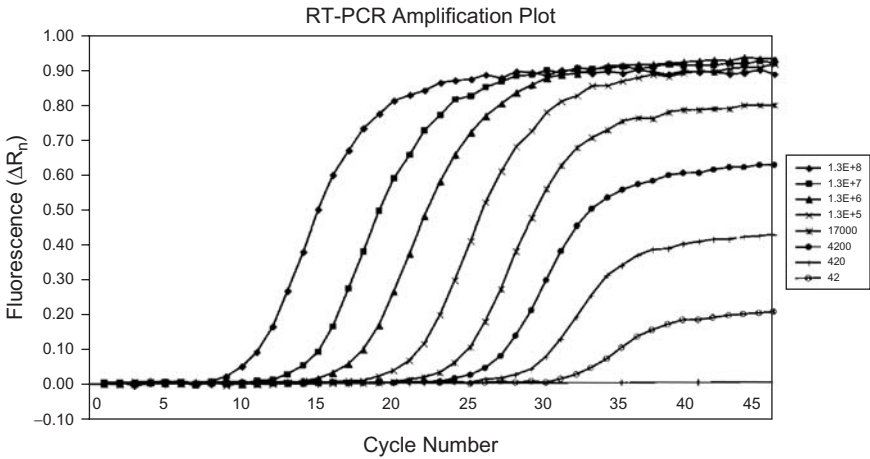


Figure 19.3 Real-time PCR measures the amount of PCR product at the end of each cycle, which is then plotted as a function of cycle number. The important parameter is the cycle threshold, or C_T , where the product mass accumulates to predetermined level. The lower the amount of starting material (genomic DNA or cDNA) the higher will be C_T . Courtesy of Dr. K.J. Livak, Applied Biosystems, Inc.

plateau effects (non-exponential product accumulation) do not compromise the quantitiveness of the assay.

There are two ways to approach data analysis in real-time PCR applications. The first is the relative quantification method, whereby the amplification of an experimental template is compared to that of a control sample. This is similar to the older, well-known relative abundance concept. The relative quantification method is also known as the $\Delta\Delta C_T$ method because one compares the C_T values of unknown samples to the amplification behavior of either housekeeping genes or the same gene in a control or reference sample. The other method, known as absolute quantification, which is a misnomer, is used to calculate the actual copy number of template molecules with statistical confidence. It involves comparing the amplification behavior of an experimental sample against a standard curve and, logically, is sometimes referred to as the standard curve method (Fig. 19.4).

Real-time systems are attractive for a number of important reasons, the most obvious of which is that data collection in real time greatly expands the linear dynamic range of this assay. The downside to implementation in the laboratory is the cost of purchasing the instrumentation, often representing an up-front investment of tens of thousands of dollars. Of course, there is also the cost of the consumables need for each experiment as well. It is important to note, however, that real-time PCR is not a suitable technique for the simultaneous assay of more than a few genes. Most often, investigators perform a preliminary screening for differentially expressed genes using microarrays, as described in Chapter 21. Following identification of purported up- or downregulated sequences, real-time PCR is then used to confirm modulation of these putative positives.

The real-time PCR format has great potential across many disciplines, including pathogen detection, molecular diagnostics, mutation detection,

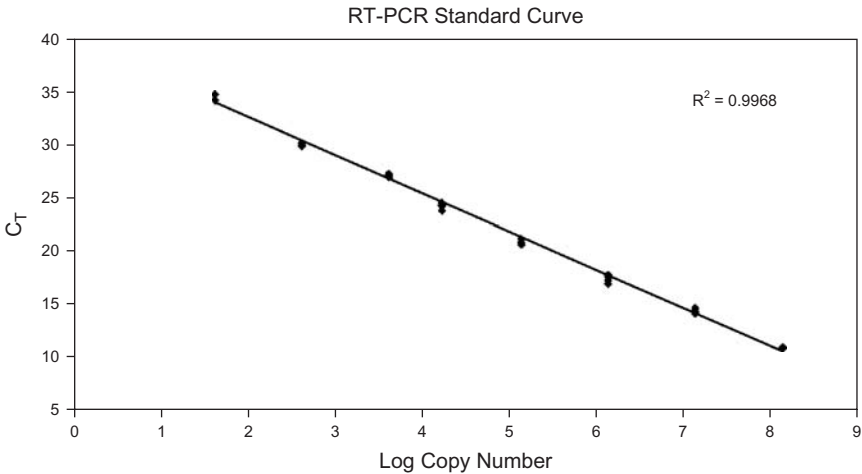


Figure 19.4 If absolute quantification (something of a misnomer) is to be performed then a standard curve is generated by the usual method of preparing and assaying serial dilutions of known mass of template, similar to the data shown in Fig. 19.3. If, however, changes in gene expression are to be reported in terms of relative abundance then the real-time format would involve the assay of a reference- or housekeeping gene (e.g., GAPDH or β -actin), or an entirely different control sample, to which the experimental gene product accumulation is compared directly. In the case of standard curve analysis, the starting number of copies of the template is plotted as a function of threshold cycle (C_T). Courtesy of Dr. K.J. Livak, Applied Biosystems, Inc.

assay of alternative splicing, the analysis of tissue samples that may have been archived long ago, for long-term in-house studies, and in high-throughput laboratories. In the case of collaborative investigations involving laboratories around the world, all of the laboratories involved must standardize their protocols and identify some type of control or reference RNA to be used for quantitative analysis. Most of the leading biotech suppliers manufacture a control RNA product that is a pool of RNA from several different cell lines representing different tissues and transcripts at different abundance levels. Because each lot of reference RNA is manufactured at the industrial scale, one may obtain RNA from the same lot for a very long time, and geographically dispersed users have ready access to the same control material.

Real-time PCR platforms

Real-time PCR combines amplification and detection, meaning that no gels are needed to analyze product. A fluorescence signal may be generated by real-time PCR in a number of interesting ways, the choice of which is based on the desired level of sensitivity as well as the nature of the data that is sought. Laboratory economics is also a concern because the instrumentation and consumables are expensive. Each of the standard platforms is described here, though it is important to realize that there are also permutations within each of these real-time methodologies.

SYBR green assay

The first published reports of the real-time PCR process actually described the use of ethidium bromide in the reaction. This intercalating agent associated itself with double-stranded PCR products as they accumulated and the resulting increase in fluorescence was proportional to the product mass. This is known as intercalator-based detection. While somewhat more reliable than many of the classical methods for PCR-based assay of transcript abundance, intercalating agents will associate both with specific and non-specific products generated in the reaction. Moreover, the intrinsic background fluorescence associated with ethidium bromide was a huge problem from the beginning, as anyone who has ever run an ethidium bromide-stained gel knows all too well.

The fluorescent dye approach has been fine-tuned by the replacement of ethidium bromide with SYBR Green, which binds to the minor groove of DNA; this dye fluoresces only when bound to double-stranded DNA (read: PCR products). The issue of a fluorescence contribution by non-specific product (false positives) persists: any double-stranded products that are synthesized, specific or non-specific, will result in measurable level of fluorescence increase (Fig. 19.5).

The SYBR Green assay has the advantage of being cost effective relative to some of the other real-time PCR methods. Because the fluorescence signal is in proportion to the product as it accumulates, there is no reason to open the tube at the end of the assay. The commonly cited disadvantages to the SYBR Green assay include the fact that SYBR Green will base-pair to any double-stranded molecules that are produced which may, therefore, result in a lower sensitivity assay. Moreover, optimization of SYBR Green-based real-time PCR is commonly necessary, though one way to minimize tube-to-tube variations is to buy a SYBR Green master mix.

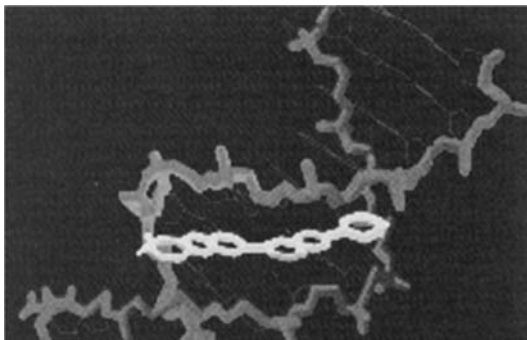


Figure 19.5 SYBR Green assay for real-time PCR. The dye binds to the minor groove of double-stranded DNA, resulting in a fluorescence signal that is proportional to the accumulating product. This methodology does not discriminate between specific or non-specific products. All double-stranded molecules will contribute to the fluorescence signal. Courtesy of Applied Biosystems, Inc.

TaqMan® assay

To address many of the intrinsic difficulties associated with intercalator-based detection and the SYBR Green assay, a modification known as probe-based detection was developed that provides a substantially higher degree of resolution. Originally known as the fluorogenic 5' nuclease assay¹ (Holland *et al.*, 1991; Lee *et al.*, 1993; Livak *et al.*, 1995), this real-time format is now known universally as the TaqMan® assay. Whereas the assay was performed originally via the liberation of a 5' ³²P end-label as an indicator and quantified by thin layer chromatography, the current form of this assay features a dual-labeled fluorochrome oligonucleotide probe, typically 18–22 bases, that is designed to anneal to the template between the region flanked by the 5' primer and the 3' primer. At its 5' end, the probe is labeled with a reporter fluorochrome, most often 6-carboxyfluorescein (6-FAM). The 3' end the probe is labeled with a non-fluorescent quencher. The spatial proximity of reporter and quencher in an intact probe suppresses fluorescence of the reporter by Förster resonance energy transfer, also known as fluorescence resonance energy transfer, or simply FRET (Förster, 1946; 1948). Because the probe is designed to have a higher T_m than the primers, the probe will anneal to the template sooner than the primers are able to anneal to their respective sites. Further, TaqMan probes are also characterized by a minor groove binder (MGB) that is linked to the quencher at the 3' end. This has the effect of increasing the T_m without increasing the length of the probe, resulting in greater sensitivity and resolution. Between 68 and 72°, the 5' exonuclease activity associated with the *Taq* polymerase displaces and then cleaves the oligonucleotide probe which is hybridized to the target strand as the upstream primer is extended. In so doing, the reporter fluorochrome is release from the probe and, no longer under the influence of the quencher, begins to fluoresce upon laser excitation. The 5'→3' polymerase activity of the *Taq* can continue through the region previously blocked by the now-displaced probe, result in the production of PCR product (Fig. 19.6). As the process is repeated, more and more fluorescence is detected via the fiberoptic system as increasing numbers of 5' reporter fluorochromes are liberated from the probe and diffuse away from the quencher. The latest generation thermal cyclers that support real-time PCR are equipped with several detection channels that will allow the investigator to perform confidently multiplex reactions. By partnering several primers, probes, and fluorochrome labels, real-time multiplex reactions greatly reduce the amount of time required to complete a set of reactions while increasing productivity dramatically.

Molecular beacons

Molecular beacon probes exhibit a characteristic stem-loop structure through which the 5' and 3' ends are maintained in close proximity (Fig. 19.7). In a manner similar to the TaqMan probes described above, fluorescence from a

¹Do not confuse this nomenclature with the nuclease protection assay (Chapter 15).

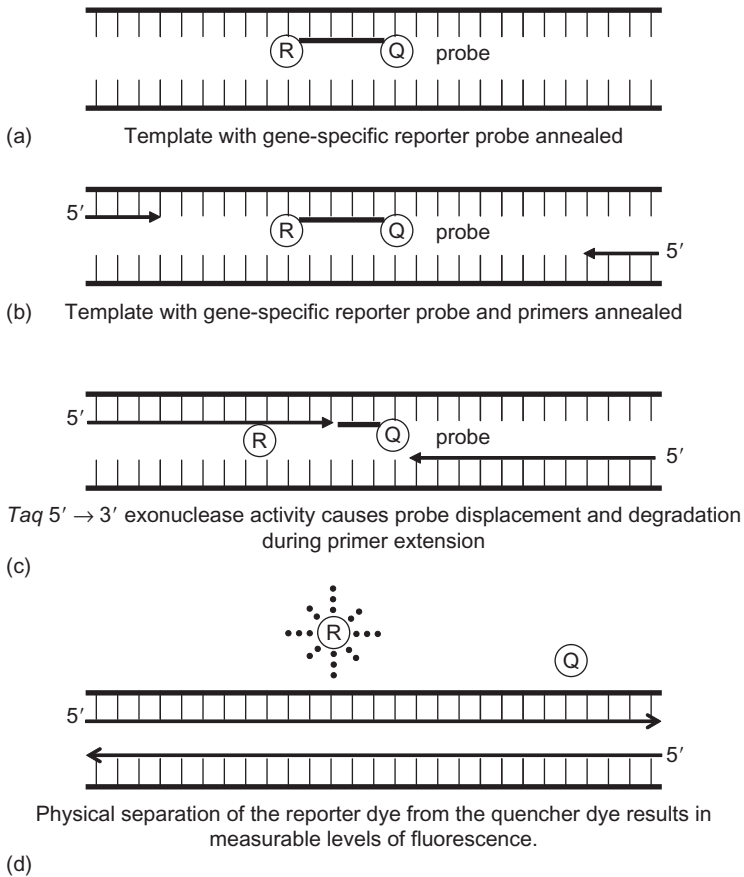


Figure 19.6 5' nuclease assay for real-time PCR. (a) Fluorogenic probes, consisting of both a fluorescent reporter dye (R) and quencher (Q), are annealed to the template. As long as probe integrity is maintained, reporter dye fluorescence is quenched due to its proximity to the quencher dye. (b) During PCR, the 5'→3' exonuclease activity of the *Taq* polymerase causes the displacement of the probe and its subsequent cleavage. (c) The fluorescence of the reporter dye, once separated from the quencher, is measured in real time (as it occurs and accumulates). Fluorescence emission is directly related to amount of PCR product being produced. PCR primer extension then proceeds through the now-exposed region, completing the product. This methodology is widely known as the TaqMan[®] assay.

fluorochrome at one end of the probe is suppressed by a nearby non-fluorescent quencher. In contrast to the TaqMan probes described above, molecular beacons are not degraded during the amplification process but rather remain intact and must bind to the target in every cycle in order to produce measurable fluorescence.

Molecular beacons are added to real-time PCR reactions prior to amplification. The disruption of base-pairing in the stem, upon denaturation, allows the

probe to hybridize to a complementary target. In so doing, the reporter dye and quencher are separated sufficiently to induce reporter dye fluorescence. As expected, simultaneous targets (multiplexing) can be detected through the use of various color fluorophores, and the technology is such that molecular beacons are able to distinguish among targets that differ by a single base, making them ideal for single nucleotide polymorphism (SNP) profiling. Molecular beacons are also quite useful in diagnostic-type assays. Licensing and other technical information pertaining to the use of molecular beacons can be accessed at www.molecular-beacons.com.

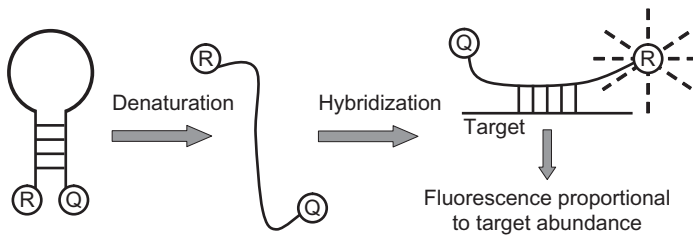


Figure 19.7 Molecular beacons. The characteristic stem-loop structure inhibits fluorescence until the beacon is denatured and able to hybridize to the accumulating product.

LUX™ primers

LUX technology (Invitrogen) is a recent innovation in the area of real-time analysis. LUX primers are labeled near the 3' end with a single fluorophore and, by virtue of their hairpin-like structure, exhibit intrinsic quenching until primer extension occurs (Fig. 19.8). The LUX nomenclature comes from “Light

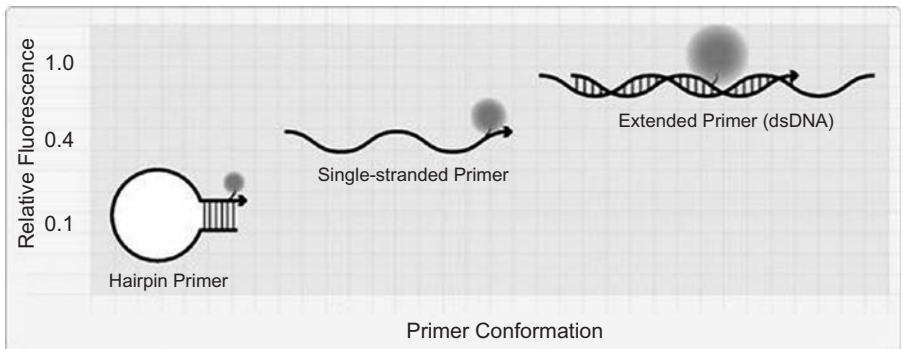


Figure 19.8 LUX technology. LUX primers do not require a quenching moiety; rather the formation of a hairpin structure naturally suppresses fluorescence until the denatured primer anneals to the template in order to support synthesis of new product by primer extension. Courtesy of Invitrogen.

Upon eXtension". LUX primers are usually 20–30 nt in length and upon incorporation into double-stranded PCR products, the relative fluorescence of the fluorophore increases approximately ten-fold. LUX primers are particularly well-suited for multiplex reactions and quantitation can be performed on any real-time instrumentation platform.

Scorpions

Scorpion probes are quite unique in that the priming and product detection are accomplished through the use of a single oligonucleotide that functions both as primer and probe (Whitcombe *et al.*, 1999). As with the TaqMan and molecular beacon probes, scorpion probes exhibit a 5' fluorophore and a downstream non-fluorescent quencher. A stem-loop structure limits fluorescence while the 3' end anneals to the template and supports PCR product synthesis. After a second high temperature denaturation, the stem-loop structure is disrupted, allowing the loop, which is complementary to the new strand that was synthesized

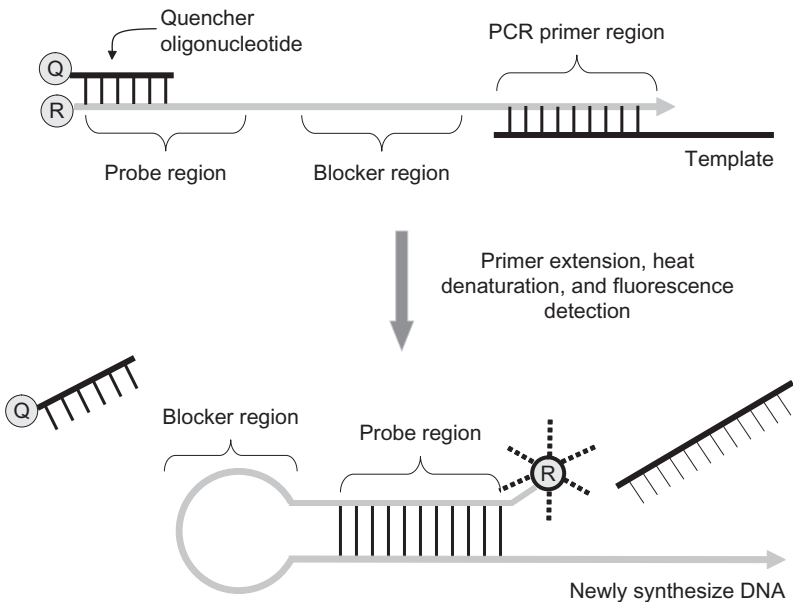


Figure 19.9 Scorpions™ Bi-probe. These molecules consist of an oligonucleotide primer linked directly to a probe and which is base-paired to another oligonucleotide with a quencher. Extension of the scorpion primer proceeds along the template to which it is base-paired. After subsequent denaturation and cooling, the scorpion probe region hybridizes in an intramolecular manner to the newly synthesized complementary DNA target. The resulting three-dimensional architecture distances the 5' fluorophore from the quencher so that fluorescence emission begins. The blocker region is chemically modified to prevent copying of the probe region. In a related design, Scorpions Uni-probes contain both the quencher and fluorophore in the same molecule (not shown).

by extension of the 3' end of the scorpion, to base pair in an intramolecular manner to the PCR product which is now part of the same molecule (Fig. 19.9). The opening of the loop and physical separation of the 5' fluorophore from the quencher results in fluorescence signal emission. This method is sometimes referred to as scorpion PCR and the resulting product might be thought of as a self-probing amplicon. Theoretically, the reaction is faster than other real-time methods because the reaction is unimolecular.

Melting curve analysis

The real-time PCR platform routinely relies on an increasing level of fluorescence as an indicator of product accumulation. Those who fear amplicon contamination feel very confident using real-time assays because the tube or capillary need not be opened post-amplification, nor is gel analysis performed. However, this begs the question: how may one determine if the correct product was produced and how can non-specific products be discerned? Needless to say, non-specific product, including primer-dimer will certainly impact sensitivity, not to mention interpretation of the data.

These questions can be answered by performing melting curve analysis (Fig. 19.10), which is a well-established method for product identification or for genotyping applications. The idea is that molecules of different length and/or base composition will dissociate or “melt apart” into their constituent single strands at different temperatures. Standard procedure is to run the melting curve protocol immediately following completion of the last cycle. As the temperature of a sample increases, double-stranded DNA will begin to melt as its T_m is approached and will complete the melting process after the T_m has been exceeded. Consequently, SYBR Green or oligonucleotides with a fluorophore label will no longer associate with single-stranded DNA, resulting in diminished fluorescence as the product(s) melt. To make the interpretation of the data easier, the melting curve profile is generally presented as the negative derivative ($-dF/dT$) so that T_m for each PCR product in the sample appears as a peak. If a unique product was produced, one should observe a single, prominent peak associated with the sample melting profile. Having *a priori* knowledge of the base composition of the expected product should also allow the investigator to predict where the melting curve peak should occur.

Problems with the reaction products are detected when the melting profile is other than expected or a second peak occurs at a lower temperature. In the case of the former, it is possible that a variant of the expected PCR product was synthesized or that a different product was produced altogether. In the case of the latter, low temperature T_m peaks often result from the spurious formation of non-specific product or, very commonly, from primer-dimer formation. In the case of a variant or otherwise unexpected product, one should examine the template, re-evaluate the primers, re-calculate the expected product size, consider raising the annealing temperature, and ensure that amplicon contamination has not occurred. Primer-dimer formation can be minimized by analyzing potential primer sequences *in silico* before ordering them. The internet is

loaded with free software that will not only examine individual primers for the likelihood of secondary structure formation and self-primer-dimer formation, but will also characterize the probable interaction of two primers that are being considered for use.

Internal controls

Observed changes in the transcriptional pattern of one gene are meaningful only when (1) compared to the behavior of the same gene in a reference (untreated or control) sample; and (2) when the transcriptional behavior of the gene of interest is compared to the transcription of other genes, specifically housekeeping genes such as GAPDH or HPRT, the expression of which are not expected to change as a function of experimental manipulation. Examples of such genes are presented in detail in Chapter 18, and the same considerations apply here as well. Thus the assay of an internal control, also known as an endogenous standard, is an indispensable component of any transcription assay. The transcripts from the housekeeping gene allow normalization (meaningful comparison) among two or more samples. Investigators should be aware, however, that there is no all-purpose control (reference) transcript applicable to every system and every circumstance.

In addition to the assay of a housekeeping gene, some investigators choose to spike an RNA sample with a known quantity of an exogenous reference template (discussed in the next section) for precision measurement of reverse transcription efficiency as well as the overall mass of PCR product at the end of the reaction. One must be aware, however, that while the spiking approach has merit with respect to accurate quantification of the expression of specific transcripts, it is not help in determining the overall transcriptional health of the cell.

Exogenous controls

The need to demonstrate differential gene expression conclusively is undisputed. For accurate calculation of transcript template copy number, rather than relative abundance, one technique is to spike an RNA sample with a known mass of an exogenous transcript. These RNA molecules can be made by *in vitro* transcription (Krieg and Melton, 1987) or simply purchased, e.g., ribulose biphosphate carboxylase (Smith *et al.*, 2003) or rabbit globin mRNA. These are examples of exogenous controls. Depending on the format of the investigation, the exogenous template can be completely unrelated to the experimental transcripts under investigation. Adding a plant transcript to animal RNA may be a particularly useful strategy in this regard: because the ribulose biphosphate carboxylase is associated with the photosynthetic pathway, no background is expected, not even basal level expression. To proceed with this type of control, add 1–5 ng of ribulose biphosphate carboxylase mRNA directly to the lysis buffer. Thus, the plant mRNA is copurified and subsequently reverse transcribed along with the RNA under investigation, thereby exposing both the

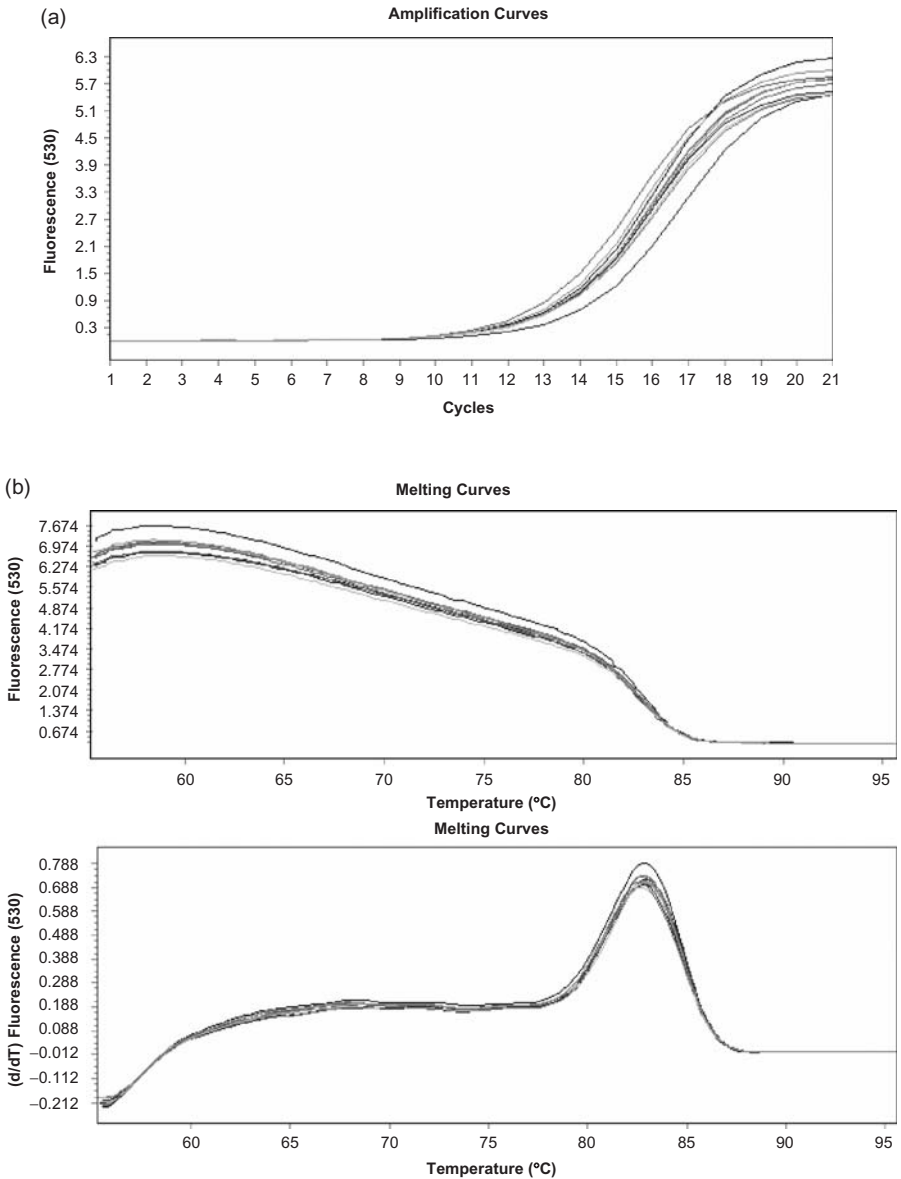


Figure 19.10 Real-time PCR and melting curve analysis. Strand dissociation causes a decrease in fluorescence as the fluorophore is released from its association with the PCR product. (a). Amplification profile of eight identical samples run in parallel. (b) Melting curve analysis of the PCR products generated in (a). The upper panel in (b) shows the fluorescence decrease as the temperature increases, though the data is most often presented as the negative derivative of the fluorescence decrease so that specific melting curve peaks can be readily identified (lower panel in b). The single peak strongly suggests a single PCR product.

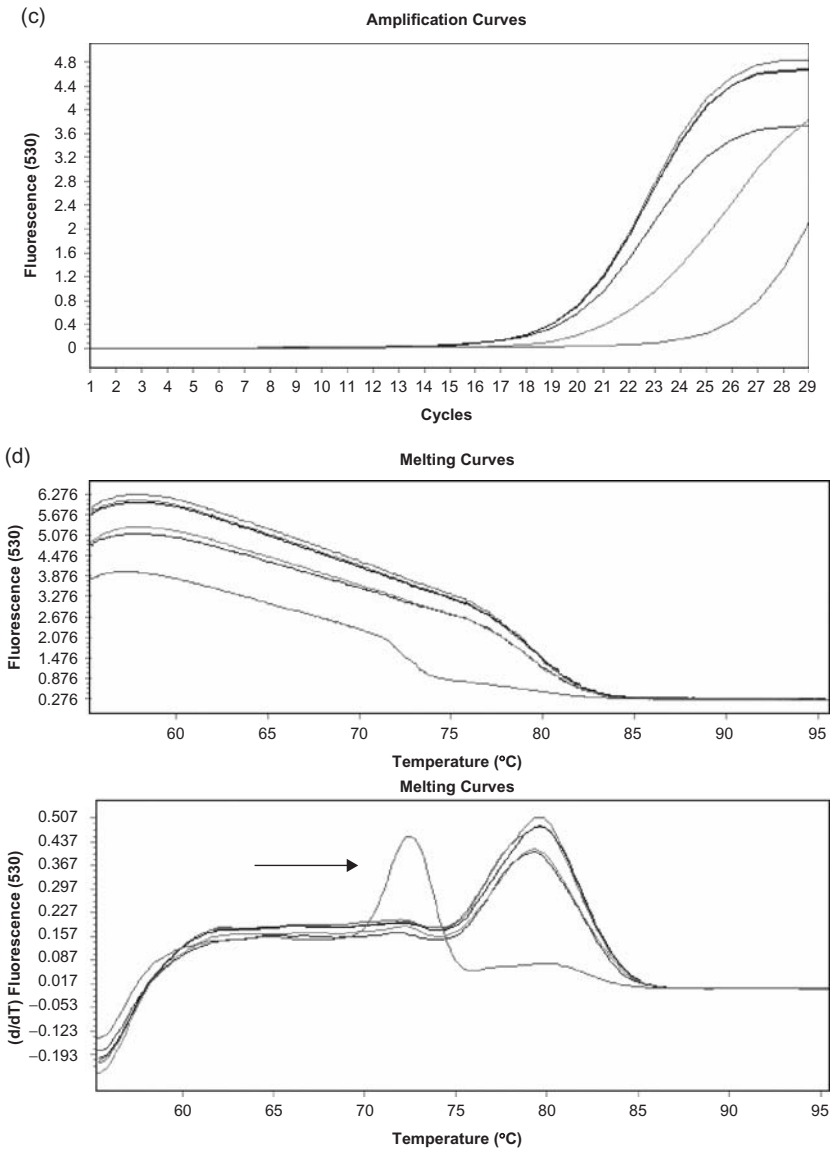


Figure 19.10 (Continued) (c). Amplification of five samples run in parallel. (d). Melting curve analysis of the PCR products generated in (c). The appearance of a second peak (arrow) at a lower than expected melting temperature strongly suggests primer-dimer or the formation of another non-specific product.

exogenous and endogenous transcripts to precisely the same chemical environments and physical manipulations during the course of the isolation procedure.

Alternatively, if the exogenous transcript is associated with at least one of the experimental transcripts to be assayed, it should be a truncated version so

that (1) it is recognized by the same primers and/or probe with the same thermodynamic stability as experimental transcript; and (2) it appears as a distinct band upon electrophoresis or as a distinct melting curve peak, thereby ensuring adequate resolution of control and experimental transcripts. Similar transcript templates such as these are sometimes referred to as “mimics”, because they mimic the characteristics of the authentic template and, as such, are co-amplified. Along these same lines, spiking an RNA sample with a bona fide transcript can be particularly useful if the RNA will also be used for RT-PCR in a downstream application. The idea is that the exogenous mRNA will be reverse transcribed and amplified by PCR with the same relative efficiency as other transcripts normally found in the sample. Having knowledge of the precise amount of globin mRNA spiked into each sample, a comparison with the amount of the corresponding PCR product can serve as a basis for normalization, as well as provide an indication of the efficiency of reverse transcription in each tube. An approach such as this is often a useful identifier of the source(s) of data variation due to inhibitors at either the cDNA synthesis- or cDNA amplification levels, particularly if there are data reproducibility issues in the lab. Further, by spiking different amounts into various samples, one is able to generate a standard curve to correlate observed experimental band intensity with the mass of the standard(s).

The notion of using an exogenous transcript as a control that would be reverse-transcribed and PCR-amplified in parallel with experimental transcripts was first described by Wang *et al.* (1989). The vector described, pAW108 (available from the American Type Culture Collection, Manassas, VA), contains sequences recognized by both 5' and 3' primers for twelve different target genes. The gene sequences are tandemly joined together, flanked by a 5' T7 RNA polymerase promoter and an encoded poly(A) tail downstream. This vector construction supports the synthesis of large quantities of complementary RNA (cRNA) which can be purified by affinity chromatography (Chapter 5) and the precise concentration determined spectrophotometrically (Chapter 6). When spiked into the experimental RNA prior to reverse transcription, the cRNA is subject to the same physical and chemical environment as the RNA under investigation, thereby minimizing concerns about the relative efficiency of reverse transcription and PCR amplification between house-keeping genes and the transcripts of interest, among various samples², and compensates for “tube effects”.

cRNA molecules can be “sense” or “antisense”, the former behaving almost identically to the native mRNA, and the latter being complementary to the native mRNA. In previous chapters, antisense transcripts were heralded as wonderful probes with the potential to form very thermodynamically stable duplexes with mRNA in applications such as the RNase protection assay. For the quantification strategy described here, however, the cRNA of interest is

²The use of cRNA standards for enhancing the quantitative dimension of Northern analysis has also been described (Wang *et al.*, 1993).

the “sense” form of the *in vitro* transcript because it must function as a minor competitor with the native mRNA during cDNA synthesis and PCR. In order to make the product of cRNA amplification distinguishable from the native RT-PCR product it is usually best to add a few bases to the *in vitro* construct or delete a few bases from the construct such that the two products, while very similar with respect to base sequence are, in fact, distinguishable. In the case of end-point analysis, another strategy is to engineer a restriction enzyme site into the template so that the resulting reference template will be exactly the same size as the experimental template and yet distinguishable from the template by post-PCR digest. In either case, the same set of primers will coamplify the two templates, the mass of each of which is related to each other by standard curve analysis.

The exogenous template methodology is widely accepted as being very sensitive, even down to the level of a single cell (Smith *et al.*, 2000) or for the assay of a single copy of a single gene (Wettwer *et al.*, 1997). The preparation and assay of cRNA standards along with experimental samples, however, can be a rather tedious process. In many applications, determining the actual mass or copy number of a transcript may not be necessary and data presentation in terms of relative abundance may be satisfactory.

Control reaction formats

Primers that are designed for both the control- and experimental transcripts can be used in a variety of formats. While these formats were developed prior to the development of real-time, PCR, if desired they can be easily adapted for real-time analysis. The following list is not intended to be exhaustive, but it does put these various historical permutations into perspective.

1. Use of a second set of primers in a different tube with an aliquot of the cDNA from the same experimental source. This is an early form of transcript quantification using PCR. In this approach the problems are severe:
 - a. The second set of primers probably will not have the same thermodynamic character as the first set and will support the synthesis of a second PCR product with a different efficiency;
 - b. Quantification in this manner requires that the mRNAs for the target and control genes be present at similar levels to permit meaningful interpretation;
 - c. mRNA for housekeeping genes often do not remain constant throughout the cell cycle;
 - d. The control and experimental samples are in different tubes, introducing the distinct possibility of “tube effects”, particularly with respect to reverse transcription efficiency.
2. Use of a second set of primers in the same PCR tube (Murphy *et al.*, 1990; Gaudette and Crain, 1991). This is also known as multiplex PCR or simply “multiplexing”. Although this is a significant improvement over the two-tube approach, the potential for serious inaccuracy persists for the same reasons cited in 1a, 1b, and 1c, above. At least both primer sets are in the same tube!

3. Competitive PCR, in which an “exogenous” internal standard, recognized by the *same* set of primers, is used. Succinctly, the cDNA under investigation *and* what is best thought of as a secondary target compete for the same primers (Becker-André and Hahlbrock, 1989; Becker-André, 1991; Gilliland *et al.*, 1990; Li *et al.*, 1991) and other reaction components. This methodology was actually described prior to the emergence of real-time PCR. In a non-real-time format, stating that this was a labor-intensive set of procedures is being generous. Samples were typically amplified over 25 cycles and the products were analyzed on agarose gels. Now, however, the amplification of competitor templates can be quantified in real-time, along with experimental templates.

There are two forms of competitive PCR. The first method involves the *in vitro* synthesis of an artificial transcript (cRNA), a precise mass of which is used to spike the experimental RNA (Wang *et al.*, 1989; Ikonen *et al.*, 1992; Vanden Heuvel *et al.*, 1993). The rationale behind the *in vitro* transcription approach is that an artificial transcript mixed directly with the experimental material will be subjected to the same physical conditions; therefore, the extent of reverse transcription is determined by assaying the amount of cDNA from the artificial transcript, whose input mass is known precisely, and using the data to construct a standard curve (Fig. 19.11). *In vitro* transcription of an artificial RNA traditionally requires the ligation of a DNA template next to a suitable transcription promoter (Fig. 19.12). Alternatively, a transcription template can be constructed by

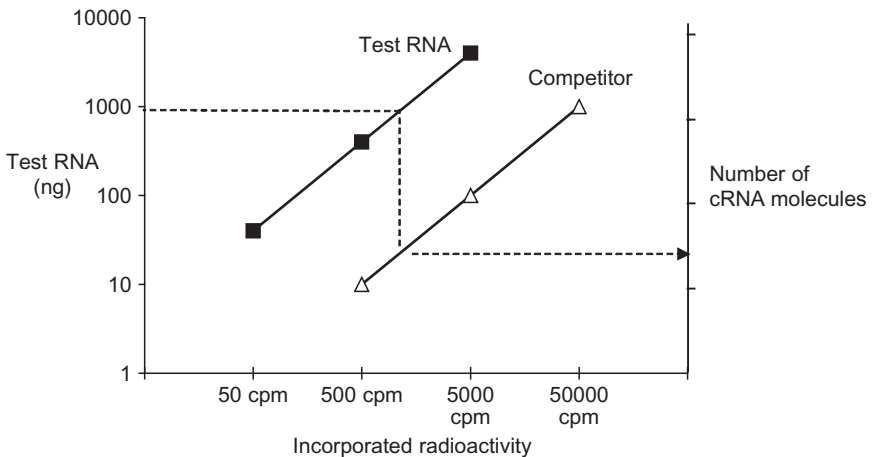


Figure 19.11 Interpolation of transcript abundance with a standard curve. In the original description of this approach (Wang *et al.*, 1989), PCR is performed using ^{32}P labeled primers. Following visualization by electrophoresis, the bands are cut out of the gel and label incorporation determined by scintillation counting. In this example ■ represents the mass of cellular RNA from a sample under investigation and Δ corresponds to the number of competitor RNA molecules used to spike the reaction tube. Equal amounts of label incorporation correlates with equal numbers of molecules of the transcript under investigation. The resulting data might then be expressed as the number of molecules of the specific transcript (determined from the coamplified cRNA) per ng or μg of RNA from the biological source.

PCR (Vanden Heuvel *et al.*, 1993) in which a bacteriophage RNA polymerase promoter is appended 5' to a DNA sequence intended for transcription (Fig. 19.13). In either case, the resulting control transcript ideally exhibits a deletion or an insertion, compared to the naturally occurring sequence. Thus, the control transcript is known as a homologous competitor because experimental and control transcripts essentially share the same sequence, including PCR primer recognition sequences; they are expected to be reverse transcribed and amplified by PCR with the same efficiency, and will be distinguishable by their size difference upon electrophoresis.

In the second approach, what might best be thought of as a non-homologous template, sharing only the primer recognition sequences with the endogenous sequence(s) under investigation, is constructed. This is accomplished by first using PCR to place primer recognition sites on the end of the competitor DNA. Composite primers, the 5' end of which is sequence (gene)-specific and the 3' end of which is competitor specific, are used to generate the competitor (Fig. 19.14). When constructed in this manner, both competitor and the corresponding cDNA will be amplified using a single set of primers. Often, non-homologous competitor sequences are used for assaying gene-specific cDNA abundance when short PCR products (less than 300bp) are expected: as the length of a template increases, so does the probability that sequence-based interference will compromise PCR efficiency. If spiked into the cDNA, rather than into the RNA, quantification in this manner does not factor in the efficiency of reverse transcription. Depending on

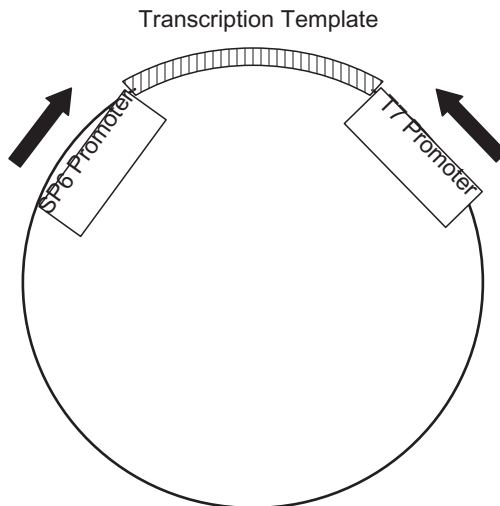


Figure 19.12 Template construction for *in vitro* transcription. The ligation of a template DNA between two different transcription promoters, such as T7 or SP6, allows the synthesis of large quantities of single-stranded sense or antisense RNA molecules from the same vector, which must be linearized prior to transcription to ensure that the resulting RNA molecules are of uniform length. The RNA can be used to spike an experimental sample as a means of determining the efficiency of reverse transcription into cDNA and subsequent amplification by PCR. Such transcripts are useful also as nucleic acid probes, details pertaining to which are found in Chapters 12 and 13.

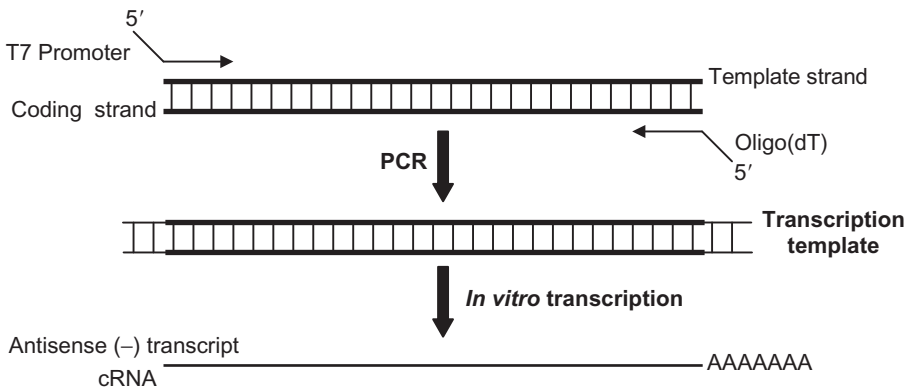


Figure 19.13 Synthesis of a transcription template by PCR. Antisense cRNA sequences are produced by *in vitro* transcription. The transcription template can be synthesized using one primer with a 5' T7 overhang and the other primer with a 5' oligo(dT) overhang. The thymidylate tract is necessary for producing an artificial poly(A) tail, which is required to support some types of reverse transcription. A similar strategy can be used to synthesize antisense cRNA molecules for use as hybridization probes.

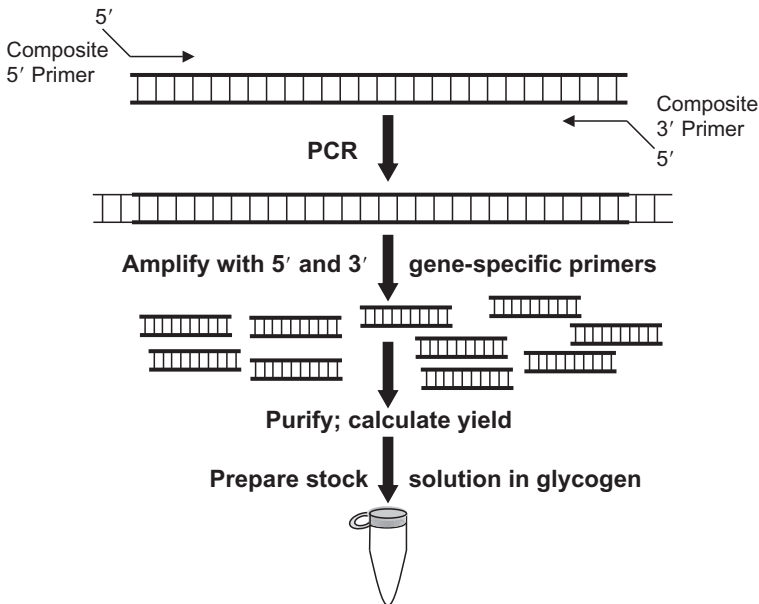


Figure 19.14 Construction of a non-homologous competitor DNA template. Both the upstream and downstream composite primers contain 5' gene-specific sequences which, following PCR, will become part of the competitor. The addition of sequence in this manner is analogous to the method of adding a restriction enzyme site, or anything else, to the ends of a PCR product. Subsequently, in the competitive PCR reactions, shorter primers, consisting only of gene-specific sequence, are used to amplify both the competitor and the cDNA under investigation. The extent of amplification of each of these species will be a direct function of their molar ratio in the reaction tube.

the context of the experiments being performed, this issue may or may not be of importance.

4. Assay a housekeeping gene, the expression of which is expected to be invariant, and relate changes in the expression of the gene of interest to it. This is particularly well-suited for real-time PCR platforms. As needed, control primers and/or the external cRNA templates described above may be included in the reaction for the generation of a standard curve (Fronhoffs *et al.*, 2002).

Negative control considerations

Negative control reactions are just as important as positive controls, quantification controls, loading controls, and enzyme efficiency controls. These reactions are designed to ensure that any observed PCR product is authentic and not an artifact due to sloppy pipetting or, worse, amplicon contamination. Many laboratories now routinely use portable PCR workstations, which look like miniature, tabletop biological safety cabinets (Fig. 19.15). These workstations typically produce HEPA³ filtered air to create a sterile pipetting environment as well as a built-in ultraviolet light that is used to crosslink rogue



Figure 19.15 Portable PCR workstation. The clean air environment with the ability to perform UV sterilization provides optimal protection from sample contamination. Courtesy of AirClean Systems.

³HEPA filter = High efficiency particulate air filter.

amplicons in the pipetting zone. Many of the more versatile workstations also feature thermoplastic construction that will protect the user from exposure to ^{32}P in experiments involving the use of isotopes.

Among the more important PCR control reactions that must be performed for each sample is the RT⁻ (RT minus) control. This consists of a tube which contains all of the components for cDNA synthesis (RNA, nucleotides, primer, etc.) *except* for reverse transcriptase. Because cDNA cannot be synthesized without reverse transcriptase, any PCR product from the RT⁻ tube is the result of contaminating genomic DNA template material. Thus, it is commonplace to treat all RNA samples with RNase-free DNase I to destroy endogenous DNA template prior to any RT-PCR assay. This has been a standard practice for RT-PCR since the inception of the technique.

More recently, it has been suggested that an RT-PCR assay can be streamlined by treating a sample with DNase I *after* the first strand synthesis. The notion is that mRNA:cDNA hybrids are resistant to DNase I so only genomic DNA, and not cDNA, will be eliminated (Flohr *et al.*, 2003). Briefly, make cDNA at 37° for 50 minutes with MMLV; heat inactivate the enzyme for 15 minutes at 50°. It is important not to exceed this temperature because dissociation of the mRNA:cDNA hybrid will render the first strand cDNA susceptible to nuclease degradation should intramolecular base-pairing occur (transient formation of a double-stranded region). The cDNA:mRNA hybrids are incubated with 10U DNase I for up to 2 hours at 37°, after which the DNase is removed by standard silica technology for cleanup. It is true that DNase I will cleave DNA:RNA and ssDNA, but does so at a much lower rate compared to dsDNA template. This approach is not for the fainthearted.

Competitive PCR: key considerations

Competitive PCR is exactly what the name implies: a form of PCR reaction in which two templates compete for amplification (Diviacco *et al.*, 1992; Siebert and Larrick, 1992; Siebert and Larrick, 1993). This is possible only if both the experimental cDNA and the competitor have identical primer recognition sites. The competitor template can be introduced either in the form of cRNA, described above, or in the form of a double-stranded cDNA after reverse transcription. Further, competitive PCR amplification can be evaluated using a real-time PCR format or by end-point analysis. The discussion which follows describes the design and synthesis of a competitor molecule which can be used in either format.

The amplification success that the experimental or competitor template will experience in a given reaction tube is a function of their molar ratios: while it might be possible to amplify both of these templates, the template represented by the larger number of molecules will yield more product, at the expense of the other template. It follows, therefore, that if both templates are present in equimolar amounts, they should be amplified with equal efficiency. Therein lies the quantitative beauty of this technique: to determine mass of an unknown

cDNA in an experimental sample, dilutions of the competitor template are made until an equimolar concentration is identified (Fig. 19.16). That's it. Plain and simple. Or is it? For the convenience of the reader the advantages and disadvantages of quantitation by standard RT-PCR and competitive PCR are compared in Tables 19.1 and 19.2.

The following are the key considerations in the design of a competitive PCR experiment:

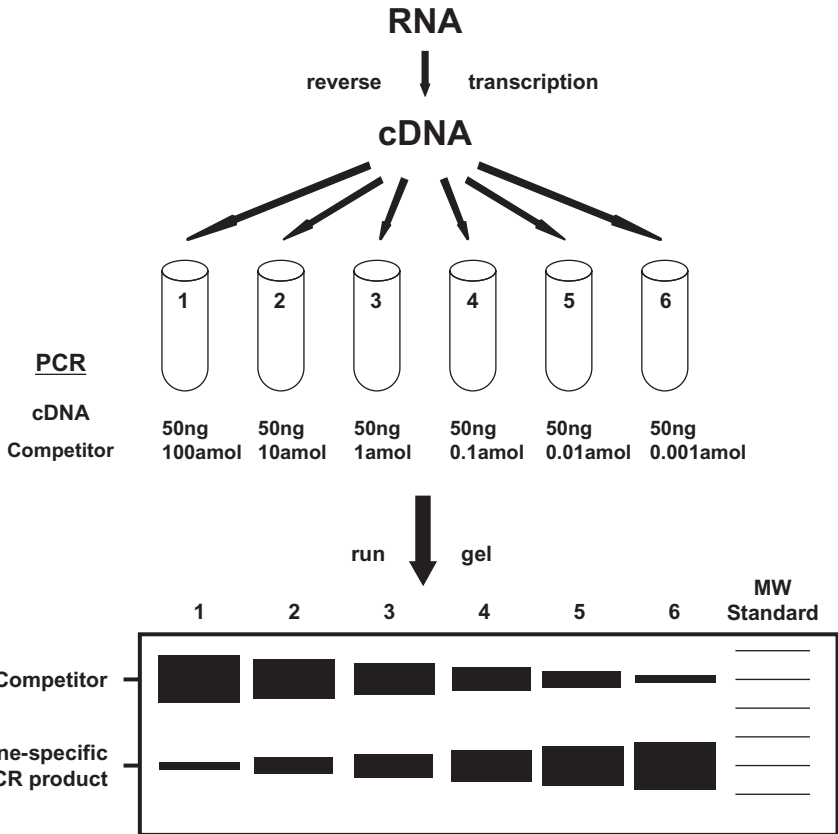


Figure 19.16 Transcript quantification by competitive PCR. Coamplification of both test cDNA and competitor, when properly designed, produces two distinct PCR products which can be discerned by electrophoresis (shown) or by melting curve analysis. In each reaction tube, both products were generated using the same set of primers and subject to the influence of the same physical parameters. The objective is to identify the reaction in which both bands are of equal intensity, meaning that the mass of each was equivalent at the onset of amplification. In cases where image analysis software is used, the abundance of the test template under investigation can be quickly determined by relating the lane where $IOD_{\text{sample}}/IOD_{\text{competitor}} = 1.0$ to the mass of competitor that was spiked into that reaction tube along with the experimental template.

1. Quantity of the competitor must be known

The competitor is synthesized by reverse transcription, PCR, or both, after which it must be purified and its concentration determined either by UV absorption (Chapter 6) or by digital image analysis (Chapter 10). Typically, the amount of competitor synthesized will be the same as for most other similar synthesis reactions, on the order of 1–3 μg . For maximum accuracy, the exact number of nucleotides that constitute the competitor must be known, from which the concentration of the newly synthesized competitor can be expressed in micromoles (μmol). When working with homologous RNA competitor transcripts, typical dilutions ranging from 1 to 10,000 femtograms (fg) of competitor are spiked into 100 ng of cellular RNA. Non-homologous competitor templates are generally diluted down to a new stock concentration in the attomole (amol) range⁴ and are often expressed in terms of actual number of molecules spiked into the tube. For quantitative analyses, the competitor is further diluted along with identical aliquots of the sample. When the amount of the experimental and competitor templates is identical, so too should be the mass of each PCR product.

Table 19.1 Standard RT-PCR Quantification: Advantages and Disadvantages

Advantages

- The format is simple.
 - The reaction is fast.
 - The reaction is versatile.
 - Very low abundance mRNAs can be detected.
 - Small numbers of cells or tissue mass are needed.
 - Poly(A)⁺ selection is neither necessary nor recommended.
 - In the case of end-point analysis, plateau-associated difficulties during amplification can be overcome by the inclusion of exogenous and/or endogenous reaction tube controls.
 - The format precludes classical difficulties of blot-associated analysis.
 - The data are reproducible.
-

Disadvantages

- Reverse transcriptase reaction is the major source of variability among samples.
 - Quantification can be difficult due to the exponential nature of PCR.
 - Plateau phase is problematic without reference transcripts.
 - Small variations in pipetting or technique can cause drastic changes in product yield and variety.
 - The reaction is sensitive to contaminants, especially genomic DNA.
 - The reproducibility of data is sometimes questionable, though this is often user-dependent.
-

⁴For the benefit of investigators unaccustomed to working with such low concentrations, a review of the units involved is presented in Table 19.3.

Table 19.2 Competitive PCR: Advantages and Disadvantages

Advantages

- One set of primers recognizes both the competitor and experimental templates.
 - cRNA control transcripts, produced by *in vitro* transcription, are useful indicators of the efficiency of reverse transcription.
 - Physical parameters in the reaction tube influence competitor and cDNA amplification equally, including plateau effects in the case of end-point PCR.
 - Quantification is based on ratios between both PCR products, so amplification need not be limited to the exponential phase.
 - The reaction is highly reproducible.
 - Products are easily distinguishable by melting curve analysis or by electrophoresis.
 - No isotope is required.
-

Disadvantages

- Several dilutions required to identify the correct ratio of target with competitor.
 - A new, unique competitor is required for each new target to be assayed.
 - A second set of primers must be synthesized to construct the competitor.
 - Non-homologous DNA competitors do not address efficiency of reverse transcription.
 - One must confirm that target and a non-homologous competitor are amplified with equivalent efficiencies before using the competitor as quantitative tool.
 - For end-point analysis, labor-intensive and time-consuming post-PCR processing is required.
-

2. Competitor and the target template amplification must be identical

One set of primers must recognize both types of template, and amplify them with the same efficiency. To accomplish this, the investigator first adds sequence which is complementary to the primers to the ends of a DNA molecule, which will become the competitor. The intended DNA competitor should show as much homology with the experimental template as possible in order to render the most accurate comparison at the end of the experiment. Thus, no one can argue that observed differences are the result of preferential amplification of one sequence over another. Very useful homologous competitor templates can be prepared by simply making an insertion or a deletion in the cDNA that is normally produced with a given set of primers. In so doing, the competitor will be either a bit smaller or a bit larger than the experimental template and will be easily resolvable electrophoretically or by melting curve analysis. In some instances when a homologous competitor sequence is not readily available, one may still confidently perform a competitive PCR experiment using a non-homologous competitor template (i.e., a sequence which is not related to the experimental cDNA under investigation) as long as:

- a. The sequence of the intended competitor is known.
- b. The intended competitor has average GC-content.

Table 19.3 Unit Analysis

Unit	Abbreviation	Mass	No. of molecules
1 mole	mol	10^0 mole	6.02×10^{23}
m, millimole	mmol	10^{-3} mole	6.02×10^{20}
μ , micromole	μ mol	10^{-6} mole	6.02×10^{17}
n, nanomole	nmol	10^{-9} mole	6.02×10^{14}
p, picomole	pmol	10^{-12} mole	6.02×10^{11}
f, femtomole	fmol	10^{-15} mole	6.02×10^8
a, attomole	amol	10^{-18} mole	6.02×10^5
z, zeptomole	zmol	10^{-21} mole	6.02×10^2

Both the stock and working concentrations of the PCR competitor are usually quite low, and investigators prefer to describe the mass of the template in attomoles. For example, a PCR product mass of 1 amol corresponds to $\approx 600,000$ molecules.

- c. The intended competitor does not have AT- or GC-rich domains.
- d. The size of the intended competitor is ± 200 bp of the size of the experimental cDNA.

3. Reverse transcription efficiency of the template and the competitor must be identical

To correlate data from any RNA-based PCR assay across a range of samples one must factor in the efficiencies of both the reverse transcriptase reaction and PCR. The coamplification of the experimental cDNA and the competitor addresses the issues surrounding the efficiency of PCR: whatever influences are exerted on the cDNA are also exerted on the competitor. Adding a DNA competitor to newly synthesized cDNA, however, does not address the issue of the relative efficiency of reverse transcription, which is always the most problematic issue surrounding cDNA synthesis. By spiking the original RNA sample(s) with a transcript produced *in vitro* prior to reverse transcription, a standard curve can be generated to address the efficiency of the reverse transcriptase in each tube. To do this, the control transcript must be designed so that it will be recognized by whatever primer is used to support reverse transcription, for example, a poly(A) tract that can be recognized by oligo(dT).

4. Heteroduplex formation must be minimized

Heteroduplexes are PCR products that form as a result of cross-hybridization (i.e., recombination) between the competitor sequence and the experimental cDNA. This is particularly common when the competitor is a deletion or insertion form of the wild-type cDNA, and these are readily identifiable as higher molecular weight products, compared to the competitor or the target cDNA alone. The greater the number of PCR cycles performed, the more likely is heteroduplex formation. Further, heteroduplex formation in the last PCR cycle is especially troublesome because these molecules will not be subsequently denatured (Ruano and Kidd, 1992; Jensen and Straus, 1993). Obviously, heteroduplex formation compromises accurate quantification of the authentic PCR products. Reducing the number of PCR cycles reduces the likelihood/severity of heteroduplex-associated difficulties.

5. Detection method must be optimized

In the case of real-time competitive PCR, the detection methodology will be melting curve analysis. However, there are many laboratories that may want to perform competitive PCR analysis for a number of reasons, in which the detection procedures are variable. Most importantly, radiolabeling the products is not required. Routinely, 25–30 cycles are performed, thereby favoring the synthesis of an adequate mass of product by simple gel staining. In this regard, SYBR Green (Schneeberger *et al.*, 1995) and SYBR Gold work well because, compared to ethidium bromide, these dyes offer a significant enhancement in sensitivity with minimal background when viewing, photodocumenting, or digitally capturing the image directly from the gel. Competitive PCR has also been described in conjunction with electrochemiluminescence (DeCesare *et al.*, 1993; Wilkinson *et al.*, 1995; Blok *et al.*, 1997). Although a reduction in the number of cycles is in order to minimize heteroduplex formation, the concomitant decrease in product accumulation need not be a detection problem because SYBR Gold in particular allows the visualization of low abundance bands to an extent not possible with ethidium bromide. Moreover, in the instance of adequate band mass, visualization with ethidium bromide is not recommended because the high background commonly associated with this dye detracts greatly from the quantitiveness of this assay. If real-time PCR is being performed then the detection is automated, as is the determination as to which cycles constitute the period of linear accumulation of product.

In summary, competitive PCR is an extremely sensitive method for template quantification. As with all PCR-based assays, the sensitivity of optimized reactions is unparalleled; data generated by Northern analysis, and even nuclease protection assay, by comparison, are lacking. The inclusion of an internal competitor that essentially mimics the experimental template precludes the requirement of having to determine in which cycles the exponential phase occurs: if the cDNA is no longer experiencing exponential amplification, neither is the competitor. Among the disadvantages associated with the method, which are really minor inconveniences more than anything else, are (1) the requirement for serial dilutions for each sample, needed to titrate the competitor template; and (2) the fact that two sets of primers must be synthesized, first a composite primer pair to synthesize the competitor, and then the gene-specific primer pair for the actual competitive PCR reactions.

Competitive PCR: major steps involved

1. Prepare PCR competitor

A. Non-homologous competitor

- Design gene-specific and composite primers (both sets of primers required). Perform the primary PCR amplification with the composite primers and then perform the secondary PCR amplification with the gene-specific primers. Try 25- to 30-mers for use as gene-specific primers, with T_m of at least 60°.

This will favor higher yields and greater fidelity of cDNA amplification. As always, use standard guidelines for PCR primer design (Chapter 18).

- Remember that competitor templates are usually 200–700 base pairs.
- Recall that competitor templates should be a maximum of 200 bp larger (or smaller) than the experimental target.
- Purify competitor product; determine concentration [spectrophotometrically or by comparison with known mass standards (e.g., Φ X174/HaeIII digest; see Chapter 8).
- Prepare titration stock solutions, usually 50–100 amol/ μ l, just prior to use, ideally in silanized (Appendix I) microfuge tubes.

B. Homologous competitor prepared by *in vitro* transcription

- Ligate cDNA adjacent to an RNA polymerase promoter (e.g., T7, SP6, or T3).
- Perform *in vitro* transcription, purify transcript, and determine concentration.
- Store at -80° until ready to use.

2. Isolate high quality RNA

- Use good RNase-free technique.
- Perform expedient recovery from cells or tissues.
- Remember that silica-based purification or a guanidinium–acid–phenol techniques are preferred.
- Treat sample(s) with RNase-free DNase.
- Wash RNA pellets with 70% ethanol prepared in nuclease-free H₂O (cationic salts influence PCR reactions by quantitatively binding of dNTPs).
- Use proper long-term storage of purified RNA (ethanol precipitate at -80°).
- If stored in aqueous buffer, aliquot and maintain at -80° .

3. Synthesize cDNA

- Select an RNase H⁻ reverse transcriptase.
- Synthesize first-strand cDNA and purify products.
- Store cDNA at -20° until ready to use.

4. Perform *primary* competitive PCR amplification (Fig. 19.17)

- Prepare serial 10-fold dilutions of the competitor.
- Use an aliquot of each PCR competitor dilution with first-strand cDNA samples.
- Perform PCR amplification; analyze products.
- Determine competitor concentration that most closely approximates target cDNA concentration.

5. Perform *secondary* competitive PCR amplification (Fig. 19.18)

- Prepare serial twofold dilutions of the competitor, beginning with a competitor stock solution that is tenfold more concentrated than the dilution that generated competitor and experimental bands of near-equal intensity in the primary amplification.
- Use an aliquot of the first-strand cDNA with each twofold competitor dilution strand.
- Perform PCR amplification and then analyze products.
- Determine competitor concentration that most closely approximates target cDNA concentration.
- Expect detection and resolution of as little as a twofold change in gene expression (transcript abundance).

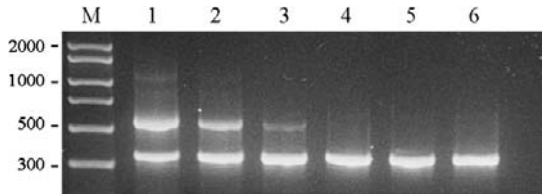


Figure 19.17 Primary tenfold dilution series for mass determination. Coamplification of mouse β -actin cDNA and a non-homologous competitor. Competitor template (540 bp) was generated using composite primers, as described. β -actin (365 bp) cDNA synthesis was supported by oligo(dT)- and AMV-mediated reverse transcription. PCR was performed for 25 cycles, after which the products of all six reactions were electrophoresed on a 2.5% agarose gel prepared in $1\times$ TAE, along with a molecular weight standard (Sigma). At the conclusion of the run, the gel was stained in $1\times$ SYBR Green diluted in $1\times$ TAE, and then photographed. The 10-fold dilution series represents a primary characterization of the sample in order to estimate the approximate mass of actin cDNA in the sample. Mass of competitor in each sample is as follows: lane 1, 10.0 amol; lane 2, 1.0 amol; lane 3, 0.1 amol; lane 4, 0.01 amol; lane 5, 0.001 amol; lane 6, 0.0001 amol. Best match ($IOD_{\text{sample}}/IOD_{\text{competitor}} = 1.0$) is in lane 2.

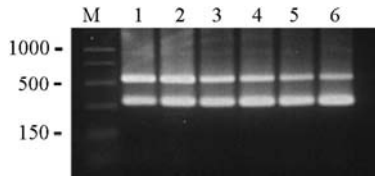


Figure 19.18 Secondary twofold dilution series for mass determination. Reaction series and analysis was identical to that described for Fig. 19.17, with the exception of the competitor mass added to each tube. Primary amplification showed the best match between competitor and experimental PCR products at 1.0 amol; thus, a twofold dilution series was utilized in the secondary amplification beginning with the previous dilution: lane 1, 10 amol; lane 2, 5 amol; lane 3, 2.5 amol; lane 4, 1.25 amol; lane 5, 0.625 amol; lane 6, 0.3125 amol. By chance, the best match is again in lane 2, meaning that the mass of actin cDNA in the sample under investigation is about 5 amol, or approximately 120,000 molecules!

Alternative approach: non-real-time competitive PCR

Real-time PCR is in widespread use, though it still beyond the means of smaller laboratories. The protocol that follows pertains to end-point analysis of competitive PCR products. It begins with the synthesis of a non-homologous competitor mediated by composite primers. The product of this reaction is then used to spike cDNA samples, thereby establishing competitive PCR amplification. If using an artificial control transcript generated by *in vitro* transcription, both the experimental RNA and the control RNA are pipetted together in the same tube and then reverse-transcribed.

Protocol: competitive PCR

Synthesis of non-homologous competitor

Each PCR competitor is synthesized in a separate reaction tube using a unique set of composite primers. Once synthesized, an aliquot of the product is used as a template for reamplification of the competitor. Note that the PCR cycling profile will vary depending on the type of thermal cycler.

1. Mix:

10× PCR buffer	5 μ l
25 mM MgCl ₂	3 μ l
10 mM dNTPs	1 μ l
Competitor template	X μ l (1 ng suggested)
10 μ M 5' composite primer	2 μ l
10 μ M 3' composite primer	2 μ l
Sterile H ₂ O	36.5–X μ l
<i>Taq</i> polymerase	0.5 μ l

2. If necessary, overlay reaction with mineral oil. Ideally, use a thermal cycler with a heated hatch and omit the mineral oil. With or without mineral oil, pulse centrifuge samples.
3. Perform cycling protocol:

Initial melt	94°	3 min	} 20 cycles
Amplification	94°	1 min	
	X°	1 min	
	72°	1 min	
Final extension	72°	10 min	
Soak	4°		

Note: Because of the extensive 5' overhang structure exhibited by each of the composite primers, it may be necessary to reduce the annealing temperature to several degrees below the calculated T_m to facilitate annealing to the competitor template.

4. Prepare a 1:500 dilution (in water) of the PCR reaction from step 3, an aliquot of which is used for reamplification. Be sure to reserve an undiluted aliquot of this same reaction as a diagnostic for electrophoresis.
5. Mix:

10× PCR buffer	5 μ l
25 mM MgCl ₂	3 μ l
10 mM dNTPs	1 μ l
Dilution from step 4	2.5 μ l
10 μ M 5' cDNA primer	2 μ l
10 μ M 3' cDNA primer	2 μ l
Sterile H ₂ O	34 μ l
<i>Taq</i> polymerase	0.5 μ l

6. If necessary, overlay reaction with mineral oil. With or without mineral oil, pulse centrifuge samples.
7. Perform cycling protocol:

Initial melt	94°	3 min	
Amplification	94°	1 min	} 20 cycles
	X°	1 min	
	72°	1 min	
Final extension	72°	10 min	
Soak	4°		

Note: If amplification is weak after 20 cycles, perform 5 more cycles.

8. Remove an aliquot of the PCR products from Step 7 and electrophorese along with the undiluted aliquot reserved in Step 4. Stain and photograph.
9. The product generated in step 8 is the competitor DNA, and it must be purified to remove the reaction components. The most efficient purification method is the use of one of the silica column-based rapid cleanup kits, such as the High Pure PCR kit (Roche).
10. Determine competitor concentration and yield, either by A_{260} or by image analysis, as described in Chapter 10. Store the competitor at -20° .
11. Just prior to use, prepare a stock solution of competitor template by diluting a portion of the competitor to a concentration of $100 \text{ amol}/\mu\text{l}$. The dilution should be made in $10 \mu\text{g}/\text{ml}$ ultrapure glycogen or in PCR TE buffer (10 mM Tris, $\text{pH } 7.5$; 0.1 mM EDTA), ideally in a silanized microfuge tube. Dilutions should be used shortly after preparation, and the remainder should be discarded.

Synthesis of first-strand cDNA

1. Ensure that high-quality RNA is available to support the synthesis of first-strand cDNA. This is ascertained by minigel electrophoresis and visualization of the ribosomal 28S and 18S RNA, with minimal smearing below the 18S rRNA. See Chapters 2–4 for isolation options.
2. Prepare the cDNA mix:

		[Final]
Nuclease-free H_2O	$3 \mu\text{l}$	–
25 mM MgCl_2	$4 \mu\text{l}$	5 mM
$10\times$ first-strand buffer	$2 \mu\text{l}$	$1\times$
10 mM dNTPs	$2 \mu\text{l}$	1 mM
Oligo(dT) ⁵	$2 \mu\text{l}$	$2.5 \mu\text{M}$
RNasin	$1 \mu\text{l}$	$2.5 \text{ U}/\mu\text{l}$

⁵Alternatively, random primers ($1 \mu\text{g}$) can be substituted for- or used in conjunction with oligo(dT) to support cDNA synthesis.

Set aside on ice.

3. Transfer 500 ng RNA (2 μ l; 250 ng/ μ l) to a sterile microfuge tube and then add 3 μ l of nuclease-free H₂O. Heat to 65° for 2 min, quickly pulse-centrifuge, and then cool on ice for 1 min.
4. Add the entire cDNA mix (from step 2) to the tube containing the heat-denatured RNA.
5. Incubate at room temperature for 3 min to facilitate primer annealing.
6. Add 1 μ l reverse transcriptase; gently pipette up and down to mix. Do not vortex, ever.
7. Incubate at 42° for 30–60 min (AMV) or at 37° for 30–60 min (MMLV).
8. Heat to 95° for 5 min to destroy the reverse transcriptase, and then store on ice until ready to use.

Competitive PCR (primary amplification)

The purpose of the primary reactions is to determine the approximate mass of the cDNA under investigation. Although the mass is expected to vary drastically among different cDNAs, reflecting the abundance of the original transcript *in vivo*, it is likely that the primary reactions will be required only once. After identifying the competitor dilution that most closely approximates the mass of the cDNA, secondary twofold dilutions of the competitor are run in order to zero in on the mass of the experimental template. Of course, profound changes in mRNA abundance, in response to an experimental challenge, will mandate repetition of the 10-fold competitor dilution series.

1. For each sample to be tested, prepare PCR master mix for 13 tubes (12 sample tubes + 1 extra) as shown in Table 19.4. Store the master mix on ice.
2. Label 6 PCR tubes P₁–P₆.
3. To prepare the primary 10-fold serial dilution series, pipette 9 μ l of 10 μ g/ml ultrapure glycogen in TE buffer (10 mM Tris, pH 7.5; 0.1 mM EDTA) into each of the 6 tubes.
4. Designate an aliquot of the competitor stock (100 amol/ μ l) as Tube P. Ten-fold dilution of the competitor is performed as follows:

Tube	Dilution method	[Final]
P ₁	add 1 μ l P, mix	10 amol/ μ l
P ₂	add 1 μ l P ₁ , mix	1.0 amol/ μ l
P ₃	add 1 μ l P ₂ , mix	10 ⁻¹ amol/ μ l
P ₄	add 1 μ l P ₃ , mix	10 ⁻² amol/ μ l
P ₅	add 1 μ l P ₄ , mix	10 ⁻³ amol/ μ l
P ₆	add 1 μ l P ₅ , mix	10 ⁻⁴ amol/ μ l

5. Label 6 new PCR tubes.
6. To each tube add:
2 μ l one of the dilutions just prepared (P₁–P₆).

48 μ l PCR master mix⁶ (step 1)

50 μ l total volume

7. Gently pulse centrifuge to collect the entire reaction volume at the bottom of the tube.
8. Perform cycling protocol.

Initial melt	94°	3 min	} 25 cycles
Amplification	94°	1 min	
	X°	1 min	
	72°	2 min	
Final extension	72°	5 min	
Soak	4°		

9. Prepare a 2.5% agarose gel, made up in 1 \times TBE or 1 \times TAE.
10. Remove a 10 μ l aliquot from each tube. If reactions were overlaid with mineral oil (old style thermal cycler), be sure to wipe the outside of the tip with a Kimwipe before transferring the aliquot into a new tube. Add 1.0 μ l 10 \times loading buffer to each sample.
11. Load gels and perform electrophoresis. Be sure to include a molecular weight standard.

Table 19.4

PCR master mix	Per 50 μ l P _x	[Final]	6 samples ($\times 7$)	12 samples ($\times 13$)
10 \times PCR buffer	5.0 μ l	1 mM	35.0 μ l	65.0 μ l
25 mM MgCl ₂ stock	3.0 μ l	1.5 mM	21.0 μ l	39.0 μ l
Sterile H ₂ O	34.5 μ l	–	241.5 μ l	448.5 μ l
10 mM dNTP mix	1.0 μ l	0.2 mM	7.0 μ l	13.0 μ l
5' primer (20 μ M)	1.0 μ l	0.4 μ M	7.0 μ l	13.0 μ l
3' primer (20 μ M)	1.0 μ l	0.4 μ M	7.0 μ l	13.0 μ l
cDNA ⁷	2.0 μ l	variable	14.0 μ l	26.0 μ l
<i>Taq</i> polymerase ⁸ (5 U/ μ l)	0.5 μ l	2.5 U/reaction	3.5 μ l	6.5 μ l
Total volume:	46.0 μ l	–	322.0 μ l	598.0 μ l

⁶Alternatively, random primers (1 μ g) can be substituted for- or used in conjunction with oligo(dT) to support cDNA synthesis.

⁷cDNA is best added to the master mix only if using the hot start PCR format because of the potential for non-specific product synthesis before the initial melt. If this occurs, then what might be thought of as an additional competitor could compromise the quantitiveness of this assay. In the absence of the hot start format, it is best to leave the cDNA out of the master mix, adding it just after the addition of the competitor in step 6.

⁸Other thermostable enzymes, especially in the hot-start format, may yield superior results compared with the use of *Taq* alone.

Note: The slower the gel is run, the better will be the resolution. Further, the concentration of the agarose can be increased or decreased as needed.

12. At the conclusion of the run, stain gel with 1× SYBR Green I diluted in 1× TBE or 1× TAE buffer. Examine gels on transilluminator (protect eyes and skin from UV light). Photodocument.
13. Determine the dilution of competitor that generated a PCR product most similar to the mass of the experimental product. While visual inspection alone is usually quite adequate for this measurement involving serial tenfold dilutions the integrated optical density (IOD) of each band and ratio for each pair of bands can be determined using image analysis software.

Competitive PCR (secondary amplification)

Having identified the competitor dilution in which the PCR products of the competitor and the cDNA are nearly identical, the competitor dilution that is tenfold *more* concentrated will be used as the starting point for making the secondary set of dilutions. The purpose for making twofold dilutions for use in a second set of reactions is to fine tune the mass determination of the gene-specific cDNA in the sample(s) under investigation. For the sake of consistency and reproducibility, use aliquots of the same PCR master mix which was prepared to support the primary reactions. The master mix should be maintained on ice; if prepared well in advance it may be frozen, as long as the *Taq* (or other) polymerase has not been added.

1. Label 6 PCR tubes S₁–S₆.
2. To prepare the twofold serial dilution series, place 5 μl of 10 μg/ml glycogen in PCR TE buffer (10 mM Tris, pH 8.0, 0.1 mM EDTA) into each of the 6 tubes.
3. Beginning with the competitor concentration that is 10-fold greater (10^x, designated as tube P_x) than that which matched the mass of the cDNA in the primary amplification, prepare the 2-fold dilution series as follows:

	Tube	Dilution method	[Final]
Dilutions:	S ₁	add 5 μl P _x , mix	5 × 10 ^{x-1} amol/μl
	S ₂	add 5 μl S ₁ , mix	2.5 × 10 ^{x-1} amol/μl
	S ₃	add 5 μl S ₂ , mix	1.25 × 10 ^{x-1} amol/μl
	S ₄	add 5 μl S ₃ , mix	6.25 × 10 ^{x-2} amol/μl
	S ₅	add 5 μl S ₄ , mix	3.125 × 10 ^{x-2} amol/μl
	S ₆	add 5 μl S ₅ , mix	1.56 × 10 ^{x-2} amol/μl

4. Label 6 new PCR tubes.
5. To each tube add:
 - 2 μl one of the dilutions just prepared (S₁–S₆).
 - 48 μl PCR master mix (see “Primary Amplification”, step 1)
 - 50 μl total volume
6. Gently pulse centrifuge to collect all reaction volume at the bottom of the tube.

7. Perform cycling protocol:

Initial melt	94°	3 min	} 25 cycles
Amplification	94°	1 min	
	X°	1 min	
	72°	2 min	
Final extension	72°	5 min	
Soak	4°		

Note: Be sure to use precisely the same cycling parameters as with the primary 10-fold dilution series, including the selection of annealing temperature.

8. Prepare a 2.5% agarose gel, made up in 1× TBE or 1× TAE. Use the same type of gel and electrophoresis parameters as for analysis of the primary reactions.
9. Remove a 10 μ l aliquot from each tube. If reactions were overlaid with mineral oil (old style thermal cycler), be sure to wipe the outside of the tip with a Kimwipe before transferring the aliquot into a new tube. Add 1.5 μ l 10× loading buffer to each sample.
10. Load gels and perform electrophoresis. Be sure to include a molecular weight standard.
11. At the conclusion of the run, stain gel with 1× SYBR Green I diluted in 1× TBE or 1× TAE buffer. Examine gels on transilluminator (protect eyes and skin from UV light). Photodocument.
12. Determine the dilution of competitor that generated a PCR product most similar to the mass of the experimental product. Because the samples represent only twofold dilutions of the competitor, it may be helpful to take advantage of image analysis software to determine the closest match between competitor and cDNA, that is, the sample in which $IOD_{\text{sample}}/IOD_{\text{competitor}} = 1$.

Image analysis considerations

The ability to make sense from the gels used to analyze competitive PCR reactions is highly dependent on the quality of the appearance and image of the gel from the start. Unfortunately, many investigators are under the sadly mistaken notion that an investment in costly image analysis systems will compensate for poorly run gels and low-quality images. The truth is that the degree to which gel analysis is accurate and reproducible is a direct function of what the gel and/or photograph look like from the onset; to a very great extent, image analysis optimization is as important as optimizing PCR or other reactions. The reader is encouraged to review the “Digital Image Analysis” section in Chapter 10 for suggestions pertaining to the optimization of image analysis.

With respect to competitive PCR, ethidium bromide-stained gels are often difficult to extract data from because of the intrinsically high levels of background associated with the use of this dye. Gels stained with SYBR Green or SYBR Gold are clearly superior. Better quality gels facilitate greater accuracy

both in visual inspection and in digital image analysis. If desired, labeled dNTPs may be included in the PCR reaction, thereby supporting detection by autoradiography and mass determinations via scintillation counting. Radiolabeling for this type of assay is not required and is often counterproductive; as such, it is not recommended.

Troubleshooting quantitative PCR techniques

The exquisite sensitivity of PCR to the precise concentrations of all reaction components is unquestioned. For assays designed to generate quantitative data, accurate pipetting and the use of master mixes for both cDNA synthesis and PCR are of critical importance. The items below outline what may be done to minimize aggravation and maximize reproducibility and productivity.

1. Non-negotiables

A. High-quality RNA:

- Use good RNase-free technique.
- Use guanidinium thiocyanate-based or guanidinium HCl-based lysis buffers.
- Perform acid-phenol extraction.
- Perform double precipitation with guanidinium lysis buffer.
- Treat samples with RNase-free DNase.
- Wash RNA pellets thoroughly and extensively with 70% ethanol prepared in nuclease-free H₂O⁹.
- Properly store purified RNA.

B. Reverse transcription into cDNA and subsequent PCR:

- Use equal mass of RNA per cDNA reaction (200–500 ng).
- Recall that enzymes lacking endogenous RNase H activity may be helpful.
- Always work with a cDNA master mix.
- Always work with a PCR master mix.

2. Common problems

A. Heavy smearing in one or more lanes:

- Optimize Mg²⁺ concentration (usually a reduction; decrease incrementally by 0.1 mM).
- Reduce the number of cycles.
- Decrease the amount of enzyme in the reaction.
- Use the hot start PCR format.

B. No product:

- Check the RNA!
- Add an RNase inhibitor (e.g., RNasin[®]).
- Test the components of the RT and PCR reactions, especially the enzymes.
- Check off each component as the reactions are assembled.
- Check PCR cycle parameters.

⁹Recall that cationic salts bind dNTPs quantitatively and can cause havoc in the cDNA synthesis and PCR reactions.

C. Low yield:

- Check for incomplete homogenization or lysis of samples.
- Check for a final RNA pellet that is incompletely redissolved.
- Determine if $A_{260}/A_{280} < 1.65$ (suggests protein or other contaminants).
- Check for RNA degradation:
 - Tissue was not immediately processed/frozen after harvesting.
 - Purified RNA was stored at -20° , instead of -80° .
 - Aqueous solutions or tubes were not RNase-free.

D. DNA contamination:

- Recall that this occurs when sample was not DNase-treated.
- Always validate with RT⁻ controls.

E. Non-reproducibility of data:

- Use thin-walled tubes (200 μ l work well); for thick walled tubes, extend cycle parameters.
- Consider hot start PCR (described in Chapter 18).
- Prepare a single, large-volume master mix per cell type and aliquot for each reaction.
- Match T_m of upstream and downstream primer pairs more closely.
- Consider enzyme mixtures (e.g., *Taq* + proofreading enzyme) designed for long-range PCR. Working together, these enzymes generate more product and do so with high fidelity.
- Consider elimination of mineral oil: use a thermal cycler with a heated lid.
- Recall that this occurs when RNA is not isolated by the same method.
- Remember that this occurs when RNA is not isolated on the same day.
- Recall that this occurs when RNA is not isolated by the same investigator.
- Ascertain the growth state of the starting cells or tissue. If cells are in different parts of the cell cycle or the tissues are isolated from biochemically different individuals, data may be non-reproducible.
- Check if a different master mix was used.
- Verify that micropipettors were calibrated recently.
- Optimize and run reactions in the same thermal cycler (not same model number).
- Consider the purchase of precast gels; both agarose and polyacrylamide are available.

3. General suggestions

- Keep denaturation steps as short as possible.
- Keep denaturation temperature as low as possible (90 – 92°).
- Reduce elongation temperature to 68° .
- Extend elongation time for each successive cycle (5 – 20 seconds per cycle).
- Hone micropipetting skills, which are critical for reproducibility. Have micropipettors recalibrated at regular intervals.
- Ensure that the input mass of RNA is exactly the same in each sample. Competitive PCR, for example, is exquisitely sensitive to the input mass of RNA (for cDNA synthesis) and cDNA (for PCR amplification). As with most RT-PCR reactions “less is more”, meaning that it is generally counterproductive to overload the reaction with too much template. When putting the PCR reactions together, it may be helpful to actually wipe the outside of the micropipette tip before dispensing the aliquot into the destination microfuge tube so that unintentional droplets on the outside of the tip do not contribute to the outcome of the reaction.

- Observe standard methods for the prevention of carryover contamination, including the use of barrier tips or positive displacement micropipettors.
- Consider the use of silanized microfuge tubes. These tubes may be helpful in minimizing losses during the various manipulations because cDNA/PCR products are present in picogram quantities (or less) prior to amplification (see Appendix I for protocol). Making dilutions of minute quantities of template in ultrapure glycogen is helpful, too.

References

- Becker-André. M. (1991). Quantitative evaluation of mRNA levels. *Methods Mol. Cell. Biol* 2, 189–201.
- Becker-André M. and K. Hahlbrock. (1989). Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcription titration assay (PATY). *Nucleic Acids Res.* 17, 9437–9446.
- Blok, H.J., A.M. Gohlke, and A.D.L. Akkermans. (1997). Quantitative Analysis of 16S rDNA using competitive PCR and the QPCR System 5000. *BioTechniques* 22, 700–704.
- DiCesare, J., B. Grossman, E. Katz, E. Picozza, R. Ragusa, and T. Woudenberg. (1993). A high-sensitivity electrochemiluminescence-based detection system for automated PCR product quantitation. *BioTechniques* 15, 152–156.
- Diviacco, S., P. Norio, L. Sentilin, S. Menzo, M. Clementi, G. Biamonti, S. Riva, A. Falaschi, and M. Giacca. (1992). A novel procedure for quantitative polymerase chain reaction by coamplification of competitive templates. *Gene* 122, 313–320.
- Flohre, A.M., T. Hackenbeck, C. Schlueter, P. Rogalla, and J. Bullerdiek. (2003). DNase I treatment of cDNA first strands prevents RT-PCR amplification of contaminating DNA sequences. *BioTechniques* 35(5), 920–926.
- Förster. T. (1946). Energiewanderung und Fluoreszenz. *Naturwissenschaften* 6, 166–175.
- Förster. T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys. (Leipzig)* 2, 55–75.
- Fronhoffs, S., G. Totzke, S. Stier, N. Wernert, M. Rothe, T. Brüning, B. Koch, A. Sachinidis, H. Vetter and Y. Ko. (2002). A method for the rapid construction of cRNA standard curves in quantitative real-time reverse transcription polymerase chain reaction. *Molecular and Cellular Probes* 16, 99–110.
- Gaudette, M.F. and W.R. Crain. (1991). A simple method for quantifying specific mRNAs in small numbers of early mouse embryos. *Nucleic Acids Res.* 19, 1879–1884.
- Gilliand, G.S., S. Perrin, K. Blanchard, and H.F. Bunn. (1990). Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction. *Proc. Natl. Acad. Sci.* 87, 2725–2729.
- Higuchi, R., G. Dollinger, P.S. Walsh and R. Griffith. (1992). Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 10, 413–417.
- Higuchi, R., C. Fockler, G. Dollinger, and R. Watson. (1993). Kinetic PCR: Real time monitoring of DNA amplification reactions. *Biotechnology* 11, 1026–1030.
- Holland, P.M., R.D. Abramson, R. Watson, and D.H. Gelfand. (1991). Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci.* 88, 7276–7280.

- Ikonen, E., T. Manninen, L. Peltonen, and A. Syvänen. (1992). Quantitative determination of rare mRNA species by PCR and solid-phase minisequencing. *PCR Methods Applic.* **1**, 234–240.
- Jensen, M.A. and N. Straus. (1993). Effect of PCR conditions on the formation of heteroduplex and single-stranded DNA products in the amplification of bacterial ribosomal DNA spacer regions. *PCR Methods Applic.* **3**, 186–194.
- Krieg, P.A. and D.A. Melton. (1987). *In vitro* synthesis with SP6 RNA polymerase. *Methods Enzymol.* **155**, 397.
- Lee, L.G., C.R. Connell, and W. Bloch. (1993). Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res.* **21**, 3761–3766.
- Li, B., P.K. Sehajpal, A. Khanna, H. Vlassara, A. Cerami, K.H. Stenzel, and M. Suthanhiran. (1991). Differential regulation of transforming growth factor beta and interleukin 2 genes in human T cells: demonstration by usage of novel competitor DNA constructs in the quantitative polymerase chain reaction. *J. Exp. Med.* **174**, 1259–1262.
- Livak, K.J., S.J. Flood, J. Marmaro, W. Guisti, and K. Deetz. (1995). Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Applic.* **4**, 357–362.
- Murphy, L.D., C.E. Herzog, J.B. Rudick, A.T. Fojo, and S.E. Bates. (1990). Use of the polymerase chain reaction in the quantification of *mdr-1* gene expression. *Biochemistry* **29**, 10351–10356.
- Ruano, G. and K.K. Kidd. (1992). Modeling of heteroduplex formation during PCR from mixtures of DNA templates. *PCR Methods Applic.* **2**, 112–116.
- Schneeberger, C., P. Speiser, F. Kury, and R. Zeillinger. (1995). Quantitative detection of reverse transcriptase-PCR products by means of a novel and sensitive DNA stain. *PCR Methods and Applic.* **4**, 234–238.
- Siebert, P.D. and J.W. Larrick. (1992). Competitive PCR (product review). *Nature* **359**, 557–558.
- Siebert, P.D. and J.W. Larrick. (1993). PCR MIMICS: Competitive DNA fragments for use as internal standards in quantitative PCR. *BioTechniques* **14**(2), 244.
- Smith, R.D., B. Brown, P. Ikonomi and A.N. Schechter. (2003). Exogenous reference RNA for normalization of real-time quantitative PCR. *BioTechniques* **34**(1), 88–91.
- Smith, R.D., J.D. Malley, and A.N. Schechter. (2000). Quantitative analysis of globin gene induction in single human erythroleukemic cells. *Nucleic Acids Res.* **28**, 4998–5004.
- Vanden Heuvel, J.P., F.L. Tyson, and D.A. Bell. (1993). Construction of recombinant RNA templates for use as internal standards in quantitative RT-PCR. *BioTechniques* **14**(3), 395–398.
- Wang, A.M., M.V. Doyle, and D.F. Mark. (1989). Quantitation of mRNA by the polymerase chain reaction. *Proc. Natl. Acad. Sci.* **86**, 9717–9721.
- Wang, S., J.J. Murtagh, Jr., C. Luo, and M. Martinez-Maldonado. (1993). Internal cRNA standards for quantitative Northern analysis. *BioTechniques* **14**(6), 935–941.
- Wettwer, C.T., M.G. Herrmann, A.A. Moss, and R.P. Rasmussen. (1997). Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques* **22**, 130–138.
- Whitcombe, D., J. Theaker, S.P. Guy, T. Brown, and S. Little. (1999). Detection of PCR products using self-probing amplicons and fluorescence. *Nature Biotechnology* **17**, 804–807.
- Wilkinson, T.E., S. Cheifetz, and S.A. DeGrandis. (1995). Development of competitive PCR and the QPCR System 5000 as a transcription-based screen. *PCR Methods Applic.* **4**, 363–367.

This page intentionally left blank

20 Functional Genomics and Transcript Profiling

Rationale 491

Functional genomics defined 491

Importance of functional genomics approaches 492

Commonly used functional genomics approaches 493

Functional genomics and classical molecular biology 496

Rationale

The completion of the human genome sequencing project, as well as the completed genome sequencing of other organisms, has rendered onto the scientific community an immense amount of data about structure and organization of the respective genomes, but very little information as to how individual genes, and the genome as a whole, actually functions.

The post-genomics era in which the scientific community finds itself has been dubbed the era of functional genomics which, depending on the source of one's information, can have rather variegated meanings. One need only perform an internet search on "functional genomics" to gain an appreciation of the lack of a strict definition – the descriptions that such a search returns are about as numerous as the number of web sites to which one is directed!

There are two fundamental goals associated with the discipline known as functional genomics. The first major goal is to survey and monitor changes in gene expression through the assay of RNA and protein abundance. While there are many useful techniques in this regard, the methods tend to be time-consuming and can become rather expensive. The second major goal is to use bioinformatics, specifically homology searches, to predict gene function from the actual DNA sequence, both within and among various species. Knowledge of this aspect of the cellular biochemistry has already made significant contributions to modern biology particularly in the areas of human health and agricultural biotechnology.

Functional genomics defined

In the mind of this Author, an appropriate way to put the notion of functional genomics into context is to think of this discipline in terms of how it relates to the

structure and organization of the genome, the study of which is known as structural genomics. For example, a good analogy is the relationship between anatomy and physiology. Anatomy is the study of biological structure while physiology is the study of biological function. The manner by which a cell, a tissue, or an organ functions is directly related to its structure. A major function of the genome is the expression (read: RNA synthesis) of the information that is encoded within it; transcription occurs as a consequence of the proper engagement of the regulatory (promoter) elements with RNA polymerase and all of the requisite ancillary factors. Whereas the genome is indicative of what a cell can potentially do, functional genomics is concerned with what a cell is actually doing at a particular moment and is, therefore, a logical extension of genome sequencing.

Another appropriate way to think about functional genomics is in terms of the methods by which genome functionality is assessed. Functional genomics tools are designed to monitor the actions of the genome from two perspectives, namely, in terms of the expression of one or a few individual genes and, more commonly, in terms of global analysis of gene expression. In particular, the ability to measure PCR product accumulation in real-time has added great value to the study of gene expression by examining RNA as one parameter of gene expression. The term ‘transcriptomics’ is also used to refer to assessing changes in the abundance of RNA in an experimental context, though ‘functional genomics’ is a more inclusive term. Finally, in order to gain a more comprehensive picture of genome functionality, one would be quite remiss if proteins were not a part of the evaluation of whole genome functionality. In the context of the dynamic nature of the transcriptome, one should include the engagement of mRNAs with the translation apparatus (translation initiation) as a logical next step after examining transcript abundance because the role of mRNA is to direct the synthesis of a polypeptide. Keep in mind, of course, that the presence of a particular mRNA in the cytoplasm does not guarantee its translation, which is another concern from a functional genomics perspective.

Gene and protein interactions are also related to functional genomics, at least tangentially. As with transcriptomics, proteomics is also a part of functional genomics that must be taken into account and which includes posttranslational modifications, protein–protein interactions, and protein turnover, not to mention the roles that proteins have in promoting transcription! Observations in this discipline lead, in turn, to metabolomics and later to systems biology. Taken together, these observations proffer a better understanding of the physiology of the cell.

Importance of functional genomics approaches

As noted above, structure directs function. In this context, it is important to understand that, while the genome is static in the sense that the semi-conservative manner in which DNA is replicated ensures genetic continuity, the genome is subject to mutations which can change the function of one or more genes. Moreover, single-nucleotide polymorphisms (SNPs) are often associated with

unique characteristics and, therefore, are interesting in a functional genomics context. Thus, qualitative data are just as important as quantitative data.

The cell uses combinatorial strategies at the DNA, RNA, and protein levels in order to create genetic diversity among the cells of a single organism. At the DNA level, for example, antibody genes and T-cell receptor genes are well known for their ability to undergo somatic recombination in lymphocytes. These are non-germ-line rearrangements that allow cells to produce a highly-specific antibody against just about any antigenic structure to which the organism is exposed. In the RNA world, posttranscriptional splicing allows the inclusion or exclusion of specific exons from a particular mRNA. Further, the selection of alternative transcription start sites can support the inclusion of a transit sequence at the amino terminus of the cognate protein, thereby allowing the protein to move into various organelles. In the protein world, immature polypeptides require a certain set of posttranslational modifications to acquire their intended function(s). In some cases, protein functionality also requires formation of a quaternary structure, meaning that two or more polypeptide subunits are required to interact precisely in order to activate the protein; the commonly used example of hemoglobin requires two α chains and two β chains in order for it function correctly in the erythrocyte.

Commonly used functional genomics approaches

This book presents a wide variety of classical and contemporary molecular biology techniques, some which are reiterated in this chapter so as to convey to the reader a sense of the variety of methods that are commonly used in the functional genomics discipline. The following list is by no means meant to be all-inclusive but rather is intended to highlight functional genomics approaches. Note that these methods are not unique to functional genomics; rather the methods are often used together in order to gain a holistic vision of genome-wide functionality.

1. **Microarray analysis** is a high-throughput method for the simultaneous assay of huge numbers of genes, as described in Chapter 21. Considered to be a global method for the analysis of gene expression, microarrays or gene chips consist of a solid support upon which very large numbers (hundreds of thousands) of individual gene sequences have been permanently applied. Newer microarray applications include the measurement of gene copy number, SNP profiling, detection of alternative RNA splicing patterns, and direct comparison of genomes from different organisms. Macroarrays, prepared by spotting cDNA onto a nylon filter, represent a less sophisticated and less expensive method for performing gene expression studies. Macroarrays are particularly useful when the entire genome of the particular organism has not been sequenced or when the investigator wishes to assay only a particular subset of genes. Thus, “themed” nylon macroarrays (e.g., cancer, apoptosis) are quite popular. Macroarrays typically contain no more than a few thousand sequences, if that, which is due in part to the porous nature of the membrane.

2. **Aptamers** are short RNA or DNA molecules which, depending on their exact nucleotide sequence, may fold into many shapes. As with many other molecules, shape confers functionality, meaning that aptamers are able to bind to specific subcellular targets and targets associated with the cell membrane for both diagnostic and therapeutic purposes – a type of smart bullet. Further, aptamers can be modified by the addition of various functional groups (Aptabodies™; www.aptagen.com) to confer drug-like properties. When aptamers that are synthesized from LNAs (locked nucleic acids) they exhibit greater stability and specificity, and are ideal for *in vivo* and *in vitro* selection. In the context of functional genomics, aptamers have great potential inasmuch as the presence and abundance of the targets to which they bind is a direct function of dynamic gene expression.
3. **Nanobiotechnology** is an area in which remarkable advances are being made for single molecule analysis and improved bioavailability of medicine, and which represent a unique blend of chemistry, biology, and physics. Nanofabrication is the science of making useful nanometer-size objects that literally involves manufacturing at the atomic level. In conjunction with microfluidic devices, also known widely as lab-on-a-chip devices, nanosensors are well-suited for the detection of extremely small quantities of DNA and RNA for disease and pathogen detection. Thus, large numbers of genes could be assayed from incredibly minute, highly differentiated cell populations or tissue samples in order to characterize genome functionality under a defined set of conditions.
4. **Gene transfer** is a process by which gene sequences are transferred into cells for over-expression or ectopic expression. Alternatively, a gene construction may be transferred into cells for the purpose of silencing another gene. Then, any of a variety of function assays may be performed to determine the consequences for the cell.
5. **PCR**, or polymerase chain reaction, is the premier method for quantitative measurement of DNA and RNA (RT-PCR). This method, described in detail in Chapters 18 and 19, revolutionized research involving all facets of the life sciences. RT-PCR has become a mainstream functional genomics tool for gene expression studies and is surely the most widely used, versatile, and arguably the most sensitive method for the assay of mRNA transcript abundance.
6. **RNAi**, or RNA interference, is a molecular biology tool for targeted gene silencing. As discussed in Chapter 23, this naturally occurring phenomenon is now sufficiently well-understood such that it can be used to target specific mRNAs for immediate destruction. In so doing, the protein synthesis is down-regulated because the template needed to support translation has been reduced if not altogether eliminated. Having revolutionized research the way PCR did in the 1990s, RNAi is considered by some to be the premier tool in the sphere of functional genomics. The science behind the technology might best be thought of as a form of reverse genetics, meaning that the function of a gene may be discovered by silencing it.
7. **SAGE**, or serial analysis of gene expression, is a nonarray method for the global analysis of gene expression (Chapter 22). Briefly, cellular RNA is converted into cDNA which is then digested into small fragments (<20 nt) known as tags. The abundance of each tag is directly related to the abundance of the corresponding mRNA from the biological source. The tags are randomly ligated together to form larger concatemers, the concatemers are sequenced, and bioinformatics methodologies are used to identify the tags and associate them with biochemistry of the original cell.
8. **Bioinformatics** represents a marriage between biotechnology and computers. As discussed in Chapter 24, bioinformatics is a result of the “information revolution”.

The idea behind bioinformatics is to be able to efficiently organize and manage the immense amounts of nucleic acid and protein sequencing data already available on-line, not to mention the fact that immense amounts of new sequence information are added to the knowledge base every day. Computers with well-written data mining software, which this Author considers to be an electronic art form, are needed to recover specific types of data and to identify patterns that are biologically meaningful with respect to gene functionality. Good places to begin bioinformatics-type queries are the National Center for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov), the Sanger Institute (www.sanger.ac.uk), and the European Bioinformatics Institute (www.ebi.ac.uk).

9. **Epigenetics** refers to changes in gene expression that are *not* the result of a change in the base sequence of a gene (i.e., changes in phenotype that are not the result of changes in genotype). Epigenetic changes may or may not be passed to descendant daughter cells. Examples of epigenetic changes include chromatin and DNA modifications such as DNA methylation, which correlates with the suppression of gene expression, and histone acetylation, which correlates with the induction of gene expression.
10. **Two-dimensional gel electrophoresis** is the classical tool for examining the proteome. Proteins are first separated by isoelectric point (pI), which is the pH at which the protein has no net charge. Then, proteins are chromatographed based on size by performing electrophoresis in a direction perpendicular to the direction of the first dimension separation. This method permits the simultaneous separation of thousands of proteins, each of which appears as a unique spot on a polyacrylamide gel; the size of each protein spot is proportional to the abundance of the protein.
11. **Protein arrays** are analogous to DNA microarrays. “Protein chips”, as they are sometimes called, consist of various proteins that have been immobilized on a glass slide or other solid support. The purpose of the array is to determine which ligands in a biological sample bind to the proteins on the slide. The process is rapid, relatively economical, and generates a lot of data from a single experiment. There are four variant protein array formats which differ in terms of what, exactly, has been immobilized on the slide, namely (a) proteins, for binding other macromolecules in a sample, including other proteins for the study of protein-protein interactions; (b) antibodies (a.k.a. antibody arrays), for the detection of specific antigens in a sample; (c) whole cell lysates representing various experimental conditions, for binding specific antibodies that recognize their cognate antigen in the lysate; and (d) affinity arrays, for the capture of any molecules to which the ligand has some type of natural biological affinity. Affinity arrays may consist of a variety of immobilized targets including, specific nucleic acid sequences, antibodies, and macromolecule complexes such as nucleic acid + protein, protein + protein, and so forth.
12. **Mass spectrometry** of simply “mass spec” is an analytical tool which, among other things, can be used to elucidate the chemical structure of various compounds, including proteins. There are several variants of this method, though the basic principal involves ionizing the compound being investigated and measuring its migration through a magnetic field, which will be a function of its charge-to-mass ratio. For example, proteins isolated from a biological source could be partially digested with trypsin, and the resulting peptide fragments analyzed by mass spec in order to discern the amino acid sequence of the original protein.
13. **Metabolomics** refers to the measurement of large sets of metabolites in cells. This is an important tool in the world of functional genomics because the induction or repression of even a single gene has the potential to result in one or more metabolic

disturbances. Metabolomics has also become quite relevant in biomarker identification. Inasmuch as metabolomics is sometimes involved with the overall metabolite composition of biological fluids, ideal matrices for biomarker development include blood and urine; changes in either of which may reflect some type of abnormal cellular function. Common metabolomics measurement methodologies include nuclear magnetic resonance (NMR), high-performance liquid chromatography (HPLC), gas-liquid chromatography (GLC), and mass spectrometry (MS).

Functional genomics and classical molecular biology

Is there really a relationship between functional genomics and classical molecular biology? The answer is a resounding “yes”! Although there is an implied association of functional genomics with all things ‘high throughput’ and ‘high tech’, even a modest molecular biology laboratory is able to perform meaningful research in this sphere with little more than standard lab equipment (thermal cycler, electrophoresis equipment, dedicated micropipettors) and access to the internet. Outsourcing DNA sequencing is commonplace and, in fact, is often more economical than in-house sequencing unless a sequencing platform is in place and is used regularly.

Keep in mind that there are methods for the global analysis of gene expression that do not involve microarrays. Suppression subtractive hybridization (SSH; Chapter 22) is a tremendously useful tool for subtracting cDNA sequences that are common to two samples, i.e., test vs. control. The remaining cDNAs represent an enriched population of uniquely expressed sequences which are then cloned and sequenced to determine the identity of each modulated gene. While SSH is not a high-throughput method, it is certainly a powerful method. One major advantage of SSH over microarray analysis is that SSH is truly a tool for new gene discovery in that any sequence that is up- or downregulated can be identified by this method. In the case of microarray analysis, one is limited to the assay of genes that have already been discovered *and* which have been printed on the array.

Similarly, mRNA differential display (Chapter 22) has an aim much like that of SSH, though the mechanics of these processes are quite a bit different. Succinctly, cDNAs representing two or more biological sources are PCR-amplified in parallel using specific primer pairs such that only a fraction of all cDNAs in a reaction tube are amplified. The judicious selection of primers ensures that only a small percentage of all cDNAs are amplified in a particular reaction which, upon electrophoresis, produces well-resolved sharp bands in each lane that profile active genes in each sample.

In addition to PCR-based techniques, including real-time- and end-point PCR, there is additional information that can be garnered through some of the time-honored molecular biology methods, including the nuclease protection assay (Chapter 15) as well as Northern analysis (Chapter 11). For example size of the native (full length) transcript can be determined by Northern analysis,

assuming that the RNA species is sufficiently abundant. Nuclease protection will provide an added level of sensitivity, though the short nature of the protected fragment compromises the ability to discern the full size of the transcript. If nothing else, these non-PCR methods are very convincing means for confirming results generated by PCR.

Classical genetics can also play a role in the realm of functional genomics. One example of such an approach involves selective breeding of organisms that exhibit some phenotype of interest. For example, mutagenesis screenings have played an important role in discerning the role of specific genes in organism-level growth and development. Briefly, random germline mutations can be fairly easily induced in population (zebrafish, for example) by exposure of some of the animals to a compound with nucleic acid mutagenizing properties. The animals exposed to this xenobiotic are then mated with wild-type zebrafish. Upon examination of the F₁, one might infer that the new phenotypes that emerge are due to xenobiotic-induced dominant mutations and are also heterozygous at the particular genetic locus. Further breeding, including a backcross with the “exposed” parent, is necessary to identify any xenobiotic-induced recessive mutations, which will be revealed in a Mendelian manner. Candidate genes responsible for the phenotype variation(s) can be identified by positional cloning, which involves classical linkage analysis, leading ultimately to DNA sequencing in order to discern the nature of the underlying mutation.

This page intentionally left blank

21 High-Throughput Analysis of Gene Expression

Rationale	499	Major steps in microarray analysis	507
What is a microarray?	500	Reference RNA	509
What is a heat map?	502	What is a macroarray?	510
What microarrays can do	504	Applications	512
What microarrays cannot do	504	References	512

Rationale

The ability to screen large numbers of samples rapidly and at the same time assay the expression of as many genes as possible has long been a desire by investigators in many disciplines. However, the primary “expression tools” including RT-PCR, nuclease protection assay, and even the relatively low sensitivity Northern analysis, valuable as they may be, assay the expression of only one or a few genes at a time and are considered low-throughput methods, no matter how fast one works in the laboratory. With the advent of microarray-based screening, large-scale expression studies are currently being performed in many laboratories and the implementation thereof has contributed a wealth of information about the function of cells and how those functions change as the cell changes. At first greeted with a fair amount of skepticism from certain elements within the scientific community, microarrays are now widely regarded both for side-by-side comparisons of two or more samples and as remarkable tools for high-throughput screening. The astonishing density of different sequences that can be placed on a microarray for simultaneous assessment qualifies microarray-based screening as a method for global (highly inclusive) analysis of gene expression. The resulting identification of the genes that are modulated under a defined set of conditions is sometimes referred to as expression profiling and is perhaps the most common microarray application. As microarray-based assays become increasingly sophisticated and more automated, productivity is expected to increase further.

What is a microarray?

A microarray, also known simply as an array (Fig. 21.1) or a gene chip, is typically a glass slide, a silicon wafer, or even a plastic substrate upon which very large numbers (currently hundreds of thousands) of individual gene sequences have been permanently applied (Fodor *et al.*, 1991; Fodor *et al.*, 1993). Microarrays were originally developed for the assay of mRNA expression from various biological sources (Schena *et al.*, 1995). Newer microarray applications include the measurement of gene copy number (Pollack *et al.*, 1999), profiling the combination of SNPs (single nucleotide polymorphisms) that characterize a sample (Hacia *et al.*, 1999), comparing the genome structure and composition of two different organisms (Moran *et al.*, 2004) or cell types, and detection of alternative splicing at exon-intron junctions (Johnson *et al.*, 2003; Ule *et al.*, 2005). It is worth noting that other microarray-based technologies are currently available such as protein microarrays (Fodor *et al.*, 1993; reviewed by Haab, 2001), commonly known as protein biochips, and antibody (Ab) microarrays. Ergo, microarray design falls into three categories: genomic arrays, which study the structure and organization of genomic DNA, transcriptomic arrays, which measure gene expression at the level of RNA synthesis, and proteomic arrays which measure protein expression and can also be used to study protein interactions. Although microarrays are considered high-throughput tools, they are not especially quantitative. Candidate genes are generally identified by high-throughput microarray analysis, after which the behavior of these genes may be more fully characterized, at least at the transcriptional level, by very quantitative real-time PCR.



Figure 21.1 A contemporary microarray – the GeneChip[®] Human Genome U133 Plus 2.0 Array. Courtesy of Affymetrix.

At present there are three nucleic acid microarray formats or platforms that are manufactured – the short oligonucleotides (25-mer) microarrays, cDNA microarrays, and the longer oligonucleotides (60- to 70-mer). Each spot that is printed onto the array typically contains picogram quantities of a unique DNA sequence. The longer oligonucleotide platform is viewed by many as the preferred format because it offers greater specificity than 25-mer oligonucleotides because of a greater tolerance for mismatches and more sequence T_m -consistency from spot to spot. Beyond 70 bases, there does not appear to be any appreciable gain in sensitivity that would justify the added cost of synthesizing longer oligonucleotides, and the likelihood of incorrect bases increases as the length of the oligonucleotide increases (reviewed by Barrett and Kawasaki, 2003). In contrast, the use of shorter oligonucleotide probes is more economical and the sequences can be applied at a much higher density. cDNA arrays often consist of high characterized segments of PCR products. Microarrays are best used within three months after purchase and should be stored at 4°.

High quality microarrays can be purchased and may cost as much as several hundred dollars each. Because of the recurrent high cost associated with the purchase of microarrays, an increasing number of well-funded labs, and institutional core facilities, are investing in the equipment needed to print their own arrays. When prepared in-house, the cost per arrays drops quickly and the arrays can be customized as needed, though quality control (QC) is often a major concern. In the case of oligonucleotide microarrays, the oligonucleotides can be synthesized directly on the surface of the microarray (e.g., photolithography) or synthesized elsewhere and then deposit onto the microarray at a later time. Cross-hybridization among homologous sequences can also be a serious problem, especially when printing microarrays with cDNAs that have been generated by PCR, which itself can be an expensive, labor-intensive effort. The gene sequences selected for inclusion on commercial microarrays are derived from several public databases that are frequently updated such as NCBI genome, RefSeq, Goldenpath, Ensembl, Unigene, RIKEN, and HUGO. The printed sequences most often include well-known, fully characterized genes, uncharacterized DNA sequences of obscure function and, if prepared in-house, custom sequences.

Glass microarrays can be used only once, while nylon macroarrays (described below) can be reused. Resolution is best, though, using fluorescence detection on glass arrays. Succinctly, microarrays represent the ability to gather qualitative and semiquantitative information about the cellular biochemistry on a mass scale, and are rapidly becoming mainstream research tools. This is not at all an unexpected development because of the enormous amount of information that has resulted from the completion of the human genome project.

In the past, each microarray was printed with sequences representing a unique tissue. Now, however, multiple-tissue microarrays are increasing in popularity, thereby facilitating the simultaneous assay of several tissues. This approach is analogous to the very popular multiple-tissue Northern blots that are sold by many biotech suppliers: RNA from several tissues has been blotted and is ready for nucleic acid hybridization. In a way, multiple-tissue

microarrays are a high-tech, high-throughput extension of *in situ* hybridization in which gene expression is assigned to specific cell types within the architecture of a tissue sample. Microarrays are also available with various themes, such as a cancer array (sometimes referred to as a cancer panel). These specialized microarrays, as well as microarrays printed with broad-ranging sequences, are designed to provide investigators with as much latitude as possible in designing their experiments.

Microarrays are probed in a very small volume (200 μ l) of hybridization buffer overnight with labeled cDNA or cRNA from two different sources, i.e., treated sample and control sample. Most often, the cDNA probes carry the fluorescent dUTPs cy3-deoxyuridine 5' triphosphate (Cy3) and cy5-deoxyuridine 5' triphosphate (Cy5), which fluoresce green and red, respectively, upon laser scanning, which is required for detection. For example, control cDNA would be labeled with Cy3 and cDNA from a treated or test sample would be labeled with Cy5. These probes are mixed together and used to co-hybridize to targets printed on the microarray (Fig. 21.2). If Cy3-labeled cDNA and Cy5-labeled cDNA hybridize to the same spot on a microarray, the laser-induced fluorescence of both Cy3 and Cy5 makes the spot appear yellow. Thus, it is the fluorescence ratio-based analysis of each spot on the microarray that provides information about the abundance of particular transcripts.

When two different cDNA populations are individually labeled with Cy5 and Cy3 and then co-hybridized to the target, as described above, the method is known as a dual-channel array. This is currently the more widely used protocol and has the advantage of requiring only one array to perform the experiment. In contrast, a one-channel microarray screening involves labeling the sample with a single fluorophore and then hybridizing that one sample to a microarray. In order to draw any conclusions, the investigator would then have to compare data generated from two or more arrays, one of which would have to be designated as the control or reference array. Succinctly, one would need to compare signal intensities from each spot on each array in order to discern up- or downregulation. At least with the dual-channel method one is able to compare two samples on single array, thereby eliminating the inevitable variability associated with two or more microarrays.

What is a heat map?

As a high-throughput method, one would naturally expect a lot of data to result from microarray work. Heat maps are one of the most useful methods for reporting data that often represent many samples, many genes, or both. A heat map is a graphical two-dimensional representation of microarray data in which the values of the data are represented in color. For example, the columns within a heat map might represent different samples while the rows might represent different genes. Regarding the color scheme, red usually indicates high expression and green usually indicates low expression.

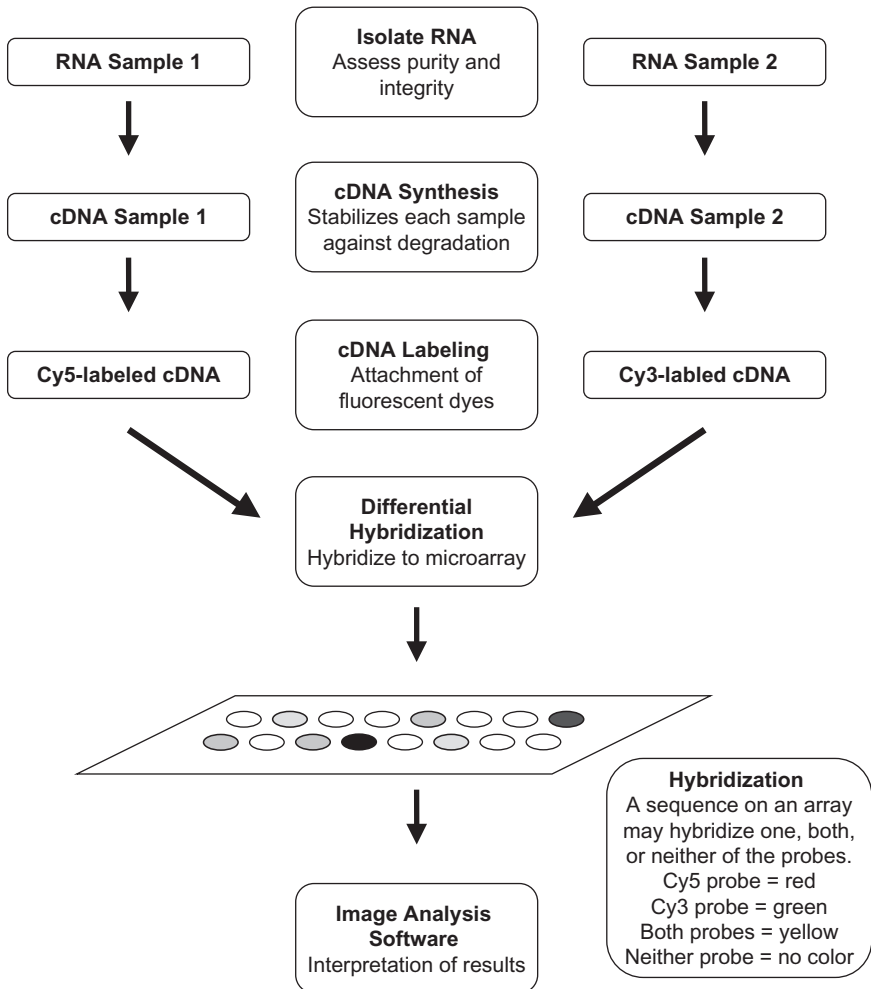


Figure 21.2 Major steps for microarray analysis. cDNAs are synthesized, labeled and hybridized to an array. Fluorescence detection or autoradiography coupled with image analysis provides a great deal of information of about patterns of gene expression in the samples under investigation.

In the construction of a heat map, it is often possible to associate specific events, for example patterns of gene expression shared by patients with a common illness, or expression of a particular gene in several tissues, organs, or organisms. The simplest way to associate gene expression similarities that become evident during the assembly of a heat map is by hierarchical clustering in which a tree-like structure known as a dendrogram groups genes showing similar levels of expression into clusters. Small clusters are then merged into progressively larger clusters until all of the genes under investigation are eventually

merged into a single cluster. Not only are genes clustered in order to render a sense of expression similarity but the clustering of individual samples is also performed. The resulting dendrogram structures are typically placed along the axes of the heat map for unambiguous interpretation of the data (Fig. 21.3).

What microarrays can do

The most obvious advantage of microarray-based profiling of gene expression is the ability to assay very large numbers of genes in a single hybridization experiment. As a result of this rapid analysis of literally thousands of genes, microarray data are now being used to implicate specific genes and gene relays in the onset and progression of various diseases. A well-planned microarray experiment asks questions about the behavior of families or groups of genes as a way of elucidating pathways associated with a specific phenotype. Microarrays provide a means for large-scale comparison of various categories of genes.

Other more traditional methods in molecular biology also offer opportunities to assay all up- and down-regulated sequences simultaneously (Table 21.1). The clear advantages with microarrays are the speed and efficiency with which huge amounts of data are acquired, processed and archived, and the immediate identification of genes in which expression is observed to fluctuate. Unlike differential display or SSH, the modulated sequences don't have to be cloned and sequenced because the sequences on the microarray have already been identified and mapped to a precise spot on the array.

How do microarrays compare to cDNA libraries? Microarray data provide a genetic snapshot of a cell or tissue undergoing any type of physiological or biochemical change, much like a collection of cDNA from a biological source and, as indicated above, the identity of microarray sequences and probable physiological significance has in most cases been defined. It is important to note that cDNA libraries are very likely to contain sequences that are not present on the microarray and which may well represent unknown genes, or mapped genes of unknown function. Microarray technology is also complementary to and compatible with most of the other contemporary tools at the disposal of the molecular biologist, including real-time PCR, RNAi, and transgenic studies. The technology is not intended to replace other time-honored molecular methods but rather to be used to embellish scientific inquiry.

What microarrays cannot do

DNA array technology has become the gold standard for the global assay of changes in gene expression. This technology, however popular, is not without intrinsic limitations.

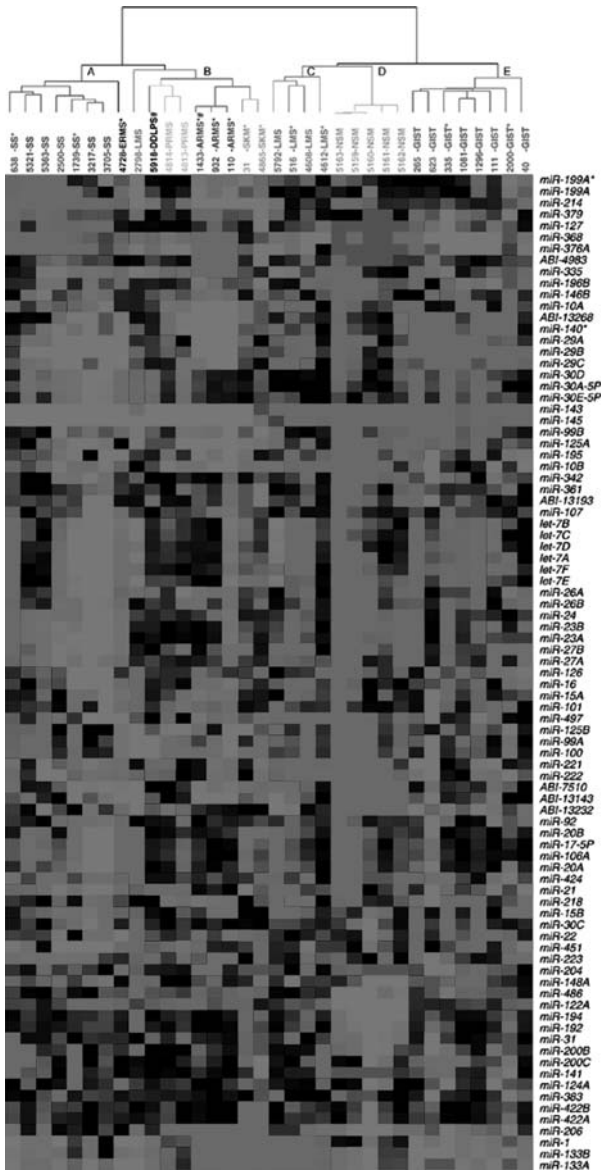


Figure 21.3 A heat map with hierarchical clustering of genes (y-axis) and samples (x-axis). Red (darker in this image) usually indicates a high level of expression, green (lighter in this image) represents a low level of expression, and the shade of red or green is an indicator of how high or how low the level of expression is. The objective behind clustering is to produce groups that behave in a similar manner. This particular image shows hierarchical clustering analysis of 27 sarcomas, 5 normal smooth muscles and 2 normal skeletal muscle tissues; each row represents the relative levels of expression for a single miRNA and each column shows the expression levels for a single sample. Reprinted by permission from Macmillan Publishers Ltd: Subramanian, S., W.O. Lui, C.H. Lee, I. Espinosa, T.O. Nielsen, M.C. Heinrich, C.L. Corless, A.Z. Fire, and M. van de Rijn. (2008). MicroRNA expression signature of human sarcomas. *Oncogene* 27, 2015–2026.

Table 21.1 Comparison of Microarray, Real-Time PCR, and Suppression Subtractive Hybridization Methods

	Microarrays	Real-time PCR	Suppression subtractive hybridization
High throughput analysis	Yes	Not the best method	No
Assay many genes simultaneously	Yes	No	Yes
Quantitativeness	Moderate	The best	Moderate
Cost	High	Moderate	Low to moderate
New gene discovery	No – can only assay sequences that are on the microarray	No – can only assay the genes for which the primers are specific	Yes! This is one of the goals of subtraction-based methods
Global analysis	Yes	No	Yes
Side-by-side comparison of two cell populations	Yes	Not easily	Yes
Reference	Universal RNA	Housekeeping genes, cRNA, or a control sample	Driver cDNA
Common probe labeling options	Cy3, Cy5	6-FAM, TAMRA (or SYBR Green)	Not applicable
Detection options	Fluorescence detection	Fluorescence detection	Gel electrophoresis
Qualitativeness	Limited	Limited	Limited
Special equipment and consumables	Requires expensive chips, microarray reader, and analysis software	Requires a real-time instrument, fluorescence-labeled probes (or SYBR Green), and analysis software	No special equipment

First, microarrays can assay only the gene sequences that have actually been printed onto the microarray. As such, microarrays are not tools for new gene discovery. The difference between microarray analysis compared to SSH and differential display is that the latter two methods render genes that differentially expressed and are not limited to the assay of a pre-defined set of genes.

Second, microarrays are semiquantitative at best. If microarray data indicate a difference in gene expression, then subsequent analysis of the gene(s) in question by real-time PCR will add an entirely new quantitative dimension

to the behavior of that gene under the conditions defined by the investigation. Microarrays are also limited in the qualitative information that they can furnish. A positive signal from a microarray gene sequence therefore indicates transcription from a particular locus, though the microarray may or may not allow recognition of alternative transcript splicing or the use of multiple transcription start sites. Elucidation of full-length transcript size is possible only by Northern analysis. Alternative start sites can be detected by nuclease protection or 5' RACE, but detection is not guaranteed by microarray analysis. Further, because the cost of microarray replicates is very high, perhaps prohibitively so for some labs, in many instances a particular sample may be assayed only once.

Third, limiting amounts of RNA starting material may actually preclude profiling a sample by microarray analysis. This is true especially when only a few cells are available, such as in laser capture, cell sorting, or when dealing with microscopically dissect tissues. In cases such as these, methods are often utilized to amplify the RNA in a linear manner¹ before attempting probe synthesis. For example, RNA may be converted into cDNA using downstream primers that incorporate a T7 or an SP6 promoter. This will allow the cDNA to be transcribed *in vitro*, thereby producing much more RNA (cRNA) that can be labeled for use directly as a probe; otherwise, the newly transcribed RNA can be converted into more cDNA that can be labeled. This strategy will not distort the natural abundance relationships among the various transcripts in the sample.

Fourth, and somewhat ironically, some of the occasional ambiguity pertaining to data analysis is a result of the fact that microarrays produce so much data all at once. As such, direct side-by-side comparisons are not always cut and dry, and most users agree that microarray data are not perfect. Nor are microarray experiments immune from interference due to contaminants and poor technique in the laboratory. The major sources of variability among experiments are (1) the biological source itself; (2) the quality of the RNA at the onset; and (3) user-dependent phenomena. Non-reproducibility problems can be further compounded by *Mycoplasma* contamination, which has been the bane of cell biologists for years and has been shown to alter microarray data when RNA is isolated from a contaminated cell culture (Miller *et al.*, 2003).

Major steps in microarray analysis

The basic steps performed in a typical microarray analysis are delineated below. Keep in mind that the method of probe labeling, method of detection, control

¹PCR is a process resulting in the exponential amplification of template sequences, thereby distorting the true differences among high-complexity samples under investigation. In contrast, *in vitro* transcription is a method that can achieve linear amplification of RNA by first converting the mRNA into cDNA and then transcribing it again *in vitro* (Eberwine method; see Chapter 18). In so doing, the true differences in transcript abundance will be preserved, even if two or more rounds are needed to produce a suitable RNA mass for labeling.

RNA, and other parameters are subject to change due to the preferences of each laboratory.

1. **Acquisition of the microarray.** As described above, microarrays are readily available from a number of sources. Of course, an investigator may be quite limited by the fact that the organism being studied is not yet represented on commercially available microarrays, and therefore must be printed in-house.
2. **Isolation of high-quality RNA.** Methods of the isolation of high-quality RNA and the assessment thereof are described in detail in Chapters 2–6. In general, RNA intended for microarray analysis requires an A_{260}/A_{280} greater than 1.8. The amount of RNA required is protocol-dependent, ranging anywhere from as little as 100 ng to as much as 10–15 μ g.
3. **cDNA synthesis.** Due to the inherent chemical instability of RNA in general, mRNA from the sample to be profiled (i.e., the “unknown” sample) is converted into cDNA much as it is for other applications. However, reverse transcription of RNA in the presence of nucleotides (e.g., dCTP) carrying a fluorescent label has been demonstrated repeatedly to be rather inefficient and skewed in favor of Cy3 incorporation. The development of a method of indirect fluorescent labeling circumvented this problem. This method involves the use of the amine derivative of dUTP, 5-(3-aminoallyl)-2'-deoxyuridine 5'-triphosphate and is commonly known as the amino allyl method of labeling. The nucleotide cocktail supporting cDNA synthesis in this manner consists of dATP, dCTP, dGTP, amino allyl-dUTP. Comparatively speaking, incorporation of this highly reactive nucleotide by reverse transcriptase has been shown to be very efficient (Hughes *et al.*, 2001). After the sample is reverse transcribed, a fluorescent dye is attached to purified cDNA in a subsequent reaction. The labeling efficiency and degree of label incorporation are most frequently assessed by scatter plot analysis (Fig. 21.4); the tighter the scatter plot, the more likely it is that the now-labeled probe will perform in a microarray assay.
4. **Linkage of fluorescent dyes.** Following the amino allyl labeling reaction, the $-NH_2$ groups incorporated into the newly synthesized cDNA are exposed to the reactive forms of Cy3 and Cy5, N-hydroxysuccinimide esters, which bind to the modified amino allyl nucleotides. This is known as the coupling reaction and is directly responsible for linking each dye to the cDNA.
5. **Purification of Cy-labeled probes.** The purpose of this very important step is to remove any unincorporated dye from the now-labeled probe. If this dye removal procedure is inefficient, the background on the microarray will be extremely high. One of the reasons why microarray data sometimes does not appear to be reproducible pertains directly to the quality of the cDNA and the overall labeling and clean-up efficiencies. High-quality cDNA comes from high-quality RNA, and careful attention must be given to the removal of unincorporated label and reaction chemistry cleanup.
6. **Microarray hybridization.** The cDNA probes are exposed directly to the array. Maintaining a minimum volume of hybridization buffer has the effect of increasing the effective probe concentration and favoring forward hybridization kinetics. Hybridization is characteristically performed under stringent conditions, followed by stringent washing to remove non-hybridized and imperfectly hybridized probe.
7. **Acquisition of data from the microarray slide or the filter.** In the case of fluorescent probes this occurs when a hybridized microarray is scanned to produce an image.

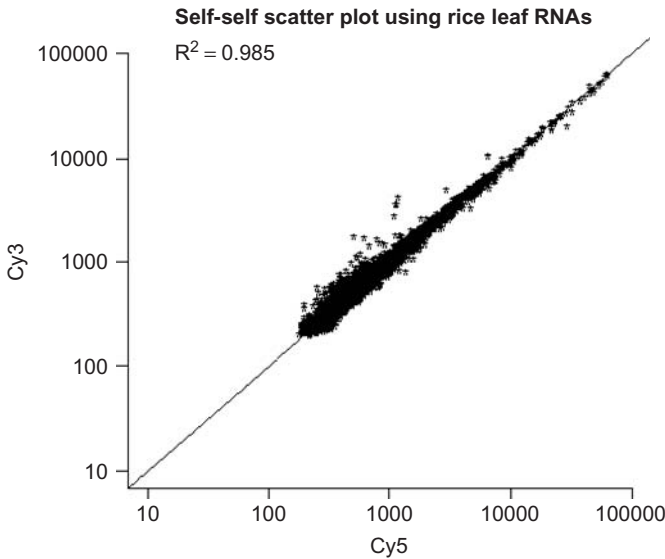


Figure 21.4 Scatter plot analysis is used to demonstrate equivalent labeling efficiencies. The fluorescence intensity from Cy3-labeled probe is plotted against the same from Cy5-labeled probe. The closer the correlation coefficient (r) is to 1.0, the better the probes will perform when hybridized to a microarray. Courtesy of Dr. Pam Ronald, University of California - Davis.

Image analysis software unique to the microarray system is then used to “manage” the image.

8. **Determination of the biological significance and relevance of data.** This is user dependent.

Reference RNA

Universal reference RNA (URR) is in widespread use in microarray experiments and is basically control RNA. In particular, it has become commonplace to use universal reference RNA in situations where a control RNA for a particular experimental model is not immediately obvious, or readily accessible, as in the study of disease progression. Universal reference RNA is very high quality RNA that is produced by pooling RNA from a number of different sources representing (1) various tissues, some of which are difficult to obtain; and (2) various abundance categories of RNA. These and other reference RNAs that are sold must pass stringent QC tests, are usually certified to have $A_{260}/A_{280} > 1.8$, are certified as being free of genomic DNA, and manifest at least a 2:1 fluorescence ratio of the 28S:18S rRNAs upon electrophoresis. QC at this level of stringency is absolutely necessary for assay optimization and, perhaps more importantly, assay standardization. A standard functionality descriptor

of universal reference RNA is known as gene coverage, which is the number of microarray sequences that hybridize to it. Gene coverage is important for comparison purposes among and between the genes under investigation. Typical coverage range is 86 to 99%.

Microarray experiments are performed using one of two basic formats:

1. **Pairwise Design.** Sample 1 and Sample 2 are labeled and hybridized onto the same microarray. For example, label sample 1 with Cy3 (green) and sample 2 with Cy5 (red) and then hybridize both to the same microarray. If there are 4 samples, then the investigator must perform all hybridizations involving all permutations:

Sample 1 + Sample 2	Sample 2 + Sample 3
Sample 1 + Sample 3	Sample 2 + Sample 4
Sample 1 + Sample 4	Sample 3 + Sample 4

for a total of 6 hybridizations. The ratio analysis performed is Sample 1/Sample 2 and, of course, must be performed all six microarrays.

2. **Universal Reference RNA Design.** This approach reduces the required number of hybridizations because each labeled sample (Cy3), in this case Samples 1–4, are individually hybridized to the microarray along with labeled (Cy5) universal reference RNA. This format requires only four hybridizations and the ratio analysis is Sample 1/URR, Sample 2/URR, Sample 3/URR, and Sample 4/URR. One of the major advantages of the universal reference design approach is that individual samples can be compared directly because each sample is compared to a single common control and allows the user to cross-reference the data to all other experiments being performed in the laboratory.² Prior to the availability of universal reference RNA researchers were not always able to compare microarray data directly due to laboratory-specific protocols. Because universal reference RNA is manufactured on an industrial scale, large numbers of aliquots representing the same lot are available for long-term studies and are simultaneously available to collaborating laboratories that are geographically dispersed.

What is a macroarray?

Multiple DNA sequences can be printed on nylon, too. In contrast to their microarray counterparts, the spots containing the DNA sequences are much larger on macroarrays. The diffusion associated with the porosity of a nylon filter limits the number of cDNAs that can be spotted on the filter to the 100s, rather than the tens- to hundreds of thousands.

Macroarrays are often themed, for example cancer or apoptosis, and are very useful for labs that have an interest in only a few genes or that do not have the

²It is strongly recommended that microarrays from the same lot be utilized when comparing closely related data sets.

equipment or resources to perform microarray analysis. Probes used to hybridize macroarrays are commonly labeled with ^{32}P and evaluated by autoradiography (Fig. 21.5). The DNA sequences on macroarrays are printed in duplicate side-by-side, so that a true positive hybridization event will produce nearly identical signals from both duplicate spots. Later, if handled properly, the hybridized probe can be removed by a high stringency wash for reuse. In this lab, we have also performed macroarray analysis using a nonradioactive hapten label in conjunction with detection by chemiluminescence with reasonably good results.

The major mechanistic difference between macroarray screening versus microarray analysis is that the detection of changes in gene expression requires a side-by-side comparison of two arrays, in which the investigator looks for increases and decreases in hybridization signal intensity. This is quite different from the fluorescence-based analysis used in conjunction with microarrays.

Macroarray advantages include (1) the geometric arrangement of sequences on the array; (2) the ability to assess the up- or downregulation of multiple gene sequences simultaneously; (3) the fact that no special equipment is necessary to determine the location and extent of hybridization on the array; (4) the ability to visually inspect the X-ray film for changes in gene expression; and (5) the relatively low cost of macroarrays.

Macroarray disadvantages are first and foremost a lack of sensitivity associated with the detection of changes in gene expression and secondly, the limited number of sequences that can be printed on a macroarray. It is more common to find protein and antibody macroarrays than nucleic acid macroarrays.

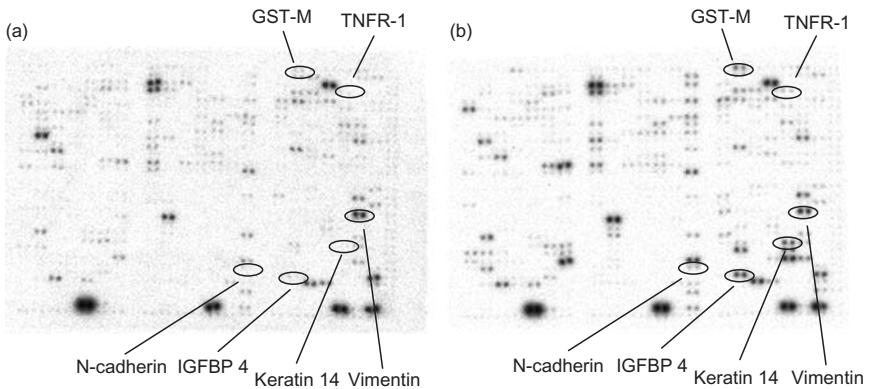


Figure 21.5 Differential expression in murine mesothelioma cell lines. The two filters represent expression in normal cells (a) and tumor cells (b). Briefly, cDNA probes were synthesized from normal and tumor mRNA in the presence of ^{32}P . Following hybridization and stringent washing, the data were generated by phosphorimager exposure and digital image analysis. Target genes on the filter are printed in pairs, and cDNA abundance is directly related to signal strength. Differential gene expression was later confirmed by Northern analysis. Courtesy of Dr. A. Kane, Brown University, Department of Pathology and Laboratory Medicine.

Applications

Commercialization of microarray technology developed from a marriage between biochemistry and the semiconductor manufacturing industry. Because of the intrinsic value of this industry to the scientific community the number of vendors has grown from one (Affymetrix) to several, including Agilent, GE Healthcare, and others. Succinctly, anyone who is studying gene expression can take advantage of microarray technology because it provides a global snapshot of gene expression, at least with respect to the sequences that are printed on the array. Unfortunately, chip replicates are expensive and, consequently, a resulting lack of data points often makes statistical analysis something of a problem. There are also upper and lower limits of detection associated with the use of microarrays. Real-time PCR is currently the only method that can very accurately quantify over a wide dynamic range changes in gene expression suggested by microarray data. Real-time PCR is also a method that validates microarray data. Microarrays reveal which genes are being modulated, and real-time PCR accurately reveals the magnitude of the modulation (by how much the expression of the gene has changed). Thus, these two methods complement each other nicely.

Clearly, microarray technology is of great interest to laboratories studying a disease process, developmental phenomena, SNP profiling, all areas of functional genomics, the response of cells to drugs, and to laboratories performing basic research and development. It is likely that microarrays in one form or another will soon be put into clinical use.

It is important to keep in mind that RNA, believe it or not, does not always tell the whole story. Gene expression is also about translation and the production of function proteins by the acquisition of the proper subset of posttranslational modifications. As such, protein chips can be used to discover which proteins are being made by the cell and what those proteins are doing. Ab microarrays constitute an analogous method for the simultaneous expression assay of hundreds of proteins, an approach known as proteome profiling. Like DNA microarrays, Ab microarrays generate data pertaining to the relative levels of a particular protein in a sample and not just its presence or absence. These microarrays consist of various monoclonal antibodies printed on the microarray in duplicate. The antibodies selected may represent a particular theme, such as a cancer panel, and the antibodies commonly used represent a numerous proteins from various locations throughout the cell. The current reported level of sensitivity of this method is about 10 pg/ml.

References

- Barrett, J.C. and E.S. Kawasaki. (2003). Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression. *Drug Discovery Today* 8, 134–141.
- Fodor, S.P., J.L. Read, M.C. Pirrung, L. Stryer, A.T. Lu, and D. Solas. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science* 251, 767–773.

- Fodor, S.P., R.P. Rava, X.C. Huang, A.C. Pease, C.P. Holmes, and C.L. Adams. (1993). Multiplexed biochemical assays with biological chips. *Nature* **364**, 555–556.
- Haab, B.B. (2001). Advances in protein microarray technology for protein expression and interaction profiling. *Curr. Opin. Drug Discov. Dev.* **4**, 116–123.
- Hacia, J.G., J.B. Fan, O. Ryder, L. Jin, K. Edgemon, G. Ghandour, *et al.* (1999). Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat. Genet.* **22**, 164–167.
- Hughes, T.R., M. Mao, A.R. Jones, J. Burchard, M.J. Marton, K.W. Shannon, *et al.* (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nature Biotech.* **19**, 342–347.
- Johnson, J.M., J. Castle, P. Garrett-Engele, Z. Kan, P.M. Loerch, C.D. Armour, *et al.* (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* **302**, 2141–2144.
- Moran, G., C. Stokes, S. Thewes, B. Hube, D.C. Coleman, and D. Sullivan. (2004). Comparative genomics using *Candida albicans* DNA microarrays reveals absence and divergence of virulence-associated genes in *Candida dubliniensis*. *Microbiology* **150**, 3363–3382.
- Miller, C.J., H.S. Kassem, S.D. Pepper, Y. Hey, T.H. Ward, and G.P. Margison. (2003). Mycoplasma infection significantly alters microarray gene expression profiles. *BioTechniques* **35**, 812–814.
- Pollack, J.R., C.M. Perou, A.A. Alizadeh, M.B. Eisen, A. Pergamenschikov and C.F. Williams. (1999). Genome-wide analysis of DNA copy number changes using cDNA microarrays. *Nat. Genet.* **23**, 41–46.
- Schena, M., D. Shalon, R.W. Davis and P.O. Brown. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.
- Subramanian, S., W.O. Lui, C.H. Lee, I. Espinosa, T.O. Nielsen, M.C. Heinrich, C.L. Corless, A.Z. Fire, and M. van de Rijn. (2008). MicroRNA expression signature of human sarcomas. *Oncogene* **27**, 2015–2026.
- Ule, J., A. Ule, J. Spencer, A. Williams, J.S. Hu, M. Cline, *et al.* (2005). Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* **37**, 796–797.

This page intentionally left blank

22 Non-Array Methods for Global Analysis of Gene Expression

Rationale	515	Troubleshooting	524
Essential issues	516	Non-subtractive methods	526
Subtractive methods	517	mRNA differential display	526
Suppression subtractive hybridization (SSH)	517	Protocols	529
		References	536

Rationale

The exquisite orchestration of gene expression at the transcriptional level and beyond is responsible for the differentiation of cells and tissues and, ultimately, for conferring the phenotype of the organism. It has long been the goal of many molecular biologists to be able to determine which genes are expressed in common among two or more different cell types and, more importantly, which genes are expressed differentially when comparing these cell populations side by side. Such insight would have profound ramifications in understanding the molecular basis of everything.

Inasmuch as the phenotype of cells is ultimately governed by the modulation of gene expression, it is logical to focus on the mRNA population of a cell and how the complexity and composition of that population changes in response to experimental challenge, compared to control or baseline conditions. Prior to the emergence of PCR, now a mainstream research tool, several variations of a method known generically as subtraction hybridization (Sargent and Dawid, 1983; Hedrick *et al.*, 1984; Fornace and Mitchell, 1986; Sargent, 1987; Orr *et al.*, 1992; Lisitsyn *et al.*, 1993; Zhang *et al.*, 1995) were utilized as a means of removing sequences shared by two transcript populations. Methods that support the physical separation of differentially expressed sequences include chromatography through hydroxyapatite¹ (Bernardi, 1971; Sargent, 1987), biotinylation coupled with streptavidin capture and modified agarose chromatography (Welcher *et al.*, 1986), PCR-based magnetic bead capture (Sharma *et al.*, 1993), biotinylation

coupled with magnetic bead separation² (Rodriguez and Chader, 1992; Lambert and Williamson, 1993; López-Fernández and del Mazo, 1993; Coche *et al.*, 1994), biotinylation coupled with phenol extraction (Travis and Sutcliffe, 1988), cross-hybridization of entire libraries (Kulesh *et al.*, 1987; Cochran *et al.*, 1987), and asymmetric PCR (Houge, 1993). Interestingly, some of these methods or variations thereof are now enjoying resurgence.

Essential issues

Regardless of the method utilized, there are terms and concepts common to all approaches that must be clearly understood if an investigation aimed at differential expression is to be well-designed. The **tester** (also known as the target) is the cDNA population containing induced sequences that are to be sequestered. The **driver** is the control or reference population of cDNA sequences against which the tester is being compared. During the subtraction hybridization, the driver is always present in a large molar excess (10-fold or more) over the tester. This will ensure first-order hybridization kinetics, resulting in complete hybridization and concomitant removal of all sequences common to both populations. It is also helpful to add fresh driver to the hybridization cocktail after several hours.

In order to understand the biochemical changes that occur within a cell in response to a particular experimental manipulation, it is quite common to perform a second subtraction in which the tester and the driver are reversed. For example, in the forward subtraction, cDNA from population A is the tester while cDNA from population B is the driver. Then, in the reverse subtraction, cDNA from population B is used as the tester and cDNA from population A becomes the driver. Performing the forward and reverse subtractions will identify induced and repressed sequences under a defined set of experimental conditions.

The overall efficiency of subtractive methods depends on maintaining high stringency conditions while favoring efficient forward hybridization kinetics, both of which are dependent upon salt concentration, hybridization temperature, hybridization time, and the presence of excess driver throughout.

After cloning, all sequences that are purportedly expressed differentially must be screened using a secondary or even tertiary method in order to confirm differential status. These methods may include but are not limited to Northern analysis (Chapter 11), nuclease protection (Chapter 15), other quantitative PCR-based methods (Chapters 18 and 19), by modified Southern analysis

¹Hydroxyapatite ($\text{Ca}_5(\text{PO}_4)_3\text{OH}$)₂ is an insoluble crystalline form of calcium phosphate. It was once used extensively for the purification of nucleic acids as well as other macromolecules. The unique nature of this matrix permits the physical partitioning of single-stranded nucleic acids from double-stranded nucleic acids (cDNA:mRNA and DNA:DNA) in a salt-dependent manner.

²See Chapter 5 for magnetic bead purification of poly(A)⁺ mRNA. The basic approach can be modified to support the purification of a number of different macromolecules by any of a number of format permutations.

(Thomas *et al.*, 1994), or by reverse Northern blot (Chapter 18). In any case, the mere presence of a sequence in a subtracted population does not guarantee the differential status of the corresponding gene.

Subtractive methods

The terms subtraction and subtraction hybridization refer to a global identification of up- and/or down-regulated sequences, as opposed to looking for modulation of specific transcripts using one or more of the techniques described in previous chapters in this text and elsewhere. Further, subtraction hybridization methods also differ dramatically from both antibody-mediated immunopurification of specific polysomes and generic polysome purification (Chapter 3). Nucleic acid-based subtraction is attractive because the end point of the assay is an enriched population of sequences partly or entirely responsible for a change in cellular phenotype. As with most contemporary RT-PCR procedures, neither RNA size selection nor poly(A)⁺ selection are requirements for efficient subtraction because of the clever design strategies for the synthesis of first-strand cDNA and amplification power of PCR. Poly(A)⁺ selection may, in fact, compromise the goals of the subtraction methodology by further under-representing rare transcripts. By linking subtraction hybridization to PCR, even very low abundance mRNAs can now be detected when they are up- or downregulated.

The ability to discriminate between expressed sequences that are shared versus those which are modulated under the conditions defined by an experiment suggests some intriguing possibilities. For example, the following are some of the comparisons that colleagues and collaborators have dreamed up over the years:

1. Quiescent cells vs. senescent cells.
2. Precursor cells vs. terminally differentiated cells.
3. Infected tissue vs. uninfected tissue.
4. Physiological changes in the kidney of certain types of fish that allow these animals to osmoregulate as they migrate from salt water into fresh water.
5. Benign tumors vs. metastatic tumors.
6. Diploid cells vs. established cell lines.
7. Apples vs. oranges.

The take-home lesson, then, is that a comparison can be made between any two populations for any reason at all, resulting the capture of modulated sequences responsible for the phenotypic differences of cells.

Suppression subtractive hybridization (SSH)

Background

The newer methods for the identification of differentially expressed genes are PCR-based approaches. In particular, the method of nucleic acid subtraction

coupled with suppression PCR (Siebert *et al.*, 1995) has become a favored method for the isolation of differentially expressed mRNAs. This approach has the advantage of being a PCR-based assay, and the unique design of the nucleic adapters needed to support this assay is the key to the system (PCR-Select™ cDNA Subtraction System; Clontech). Moreover, the mechanics of the assay (Fig. 22.1) favor a normalization of differentially expressed sequences, meaning that the end of the assay produces an enriched pool of differentially expressed sequences that are present at similar concentrations, regardless of their abundance level in the original biological material. This situation is highly desirable during the identification process because highly abundant cDNAs are more likely to be found in a non-normalized library than lesser- or underrepresented cDNAs. Thus, sequencing 100 clones from a normalized sublibrary will provide a much more comprehensive snapshot of the variety of up- or down-regulated sequences, compared to a library where the relative abundance of the sequences *in vivo* is preserved.

In this laboratory the PCR-Select method has been used in several applications, with outstanding results (Norelli *et al.*, 2009; Bassett *et al.*, 2006). While technically a bit more challenging than some of the older methods, subtraction-suppression PCR greatly enriches in favor of low abundance transcripts and, in the mind of this Author, generates a more complete profile of gene expression than differential display. One can expect that any sequence which differs in abundance by fivefold will be easily sequestered by the subtraction process, though it is not uncommon for the PCR-Select approach to render genes with as little as a 1.5-fold difference compared to the driver population.

Protocols

Succinctly, the methodology (Fig. 22.2) involves creating two tester sub-populations with different adapters for each cDNA population, while no adapters are ligated to the driver. Efficient ligation is favored by restricting the tester cDNA with the enzyme *Rsa* I, a four-base cutter that prepares blunt-ended restriction fragments. The uniqueness of this approach lies with the fact that the adapters do not exhibit 5' phosphate groups, meaning that (1) the cDNA itself will be the phosphate group donor; (2) only one strand of each adapter will be joined to each cDNA; and (3) stretches of identical sequences within the adapters will facilitate the annealing of primers, and thus PCR, after a simple filling-in reaction has been performed.

The results of two hybridizations and the appropriate mixing of the two tester sub-populations (Fig. 22.3), are templates for PCR amplification that represent the differentially expressed sequences. Then, a primary and secondary set of amplifications with nested primers renders an enriched, equalized population of subtracted sequences by virtue of the PCR suppression phenomenon that occurs (Fig. 22.4). The enriched, differentially-expressed sequences

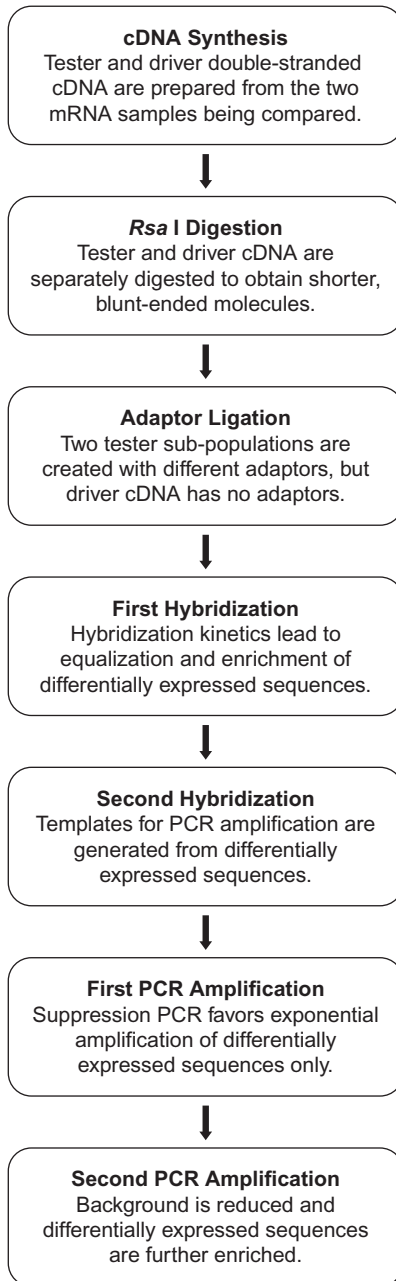


Figure 22.1 Overview of the PCR-Select cDNA Subtraction system. Courtesy of Clontech Laboratories.



Figure 22.2 Preparation of adaptor-ligated tester cDNAs for hybridization and PCR. To perform the forward and reverse subtractions, each tester cDNA must be ligated to the appropriate adaptors as shown. Courtesy of Clontech Laboratories.

are now ready for cloning, confirmation of differential status, and sequencing (Fig. 22.5). Performing the secondary PCR amplification with nested primers is absolutely essential for diluting out low-molecular-weight junk and also generating an adequate mass of larger, useful cDNAs for identification purposes. In this lab, when an attempt was made to clone the subtracted PCR products generated by the primary PCR amplification only, nothing but low-molecular-weight DNA was ligated into the vector, and not a single clone produced a meaningful insert upon colony PCR screening.

There are two hybridizations in this subtraction-suppression method, both of which must be performed correctly. The first hybridization requires 6 to 12

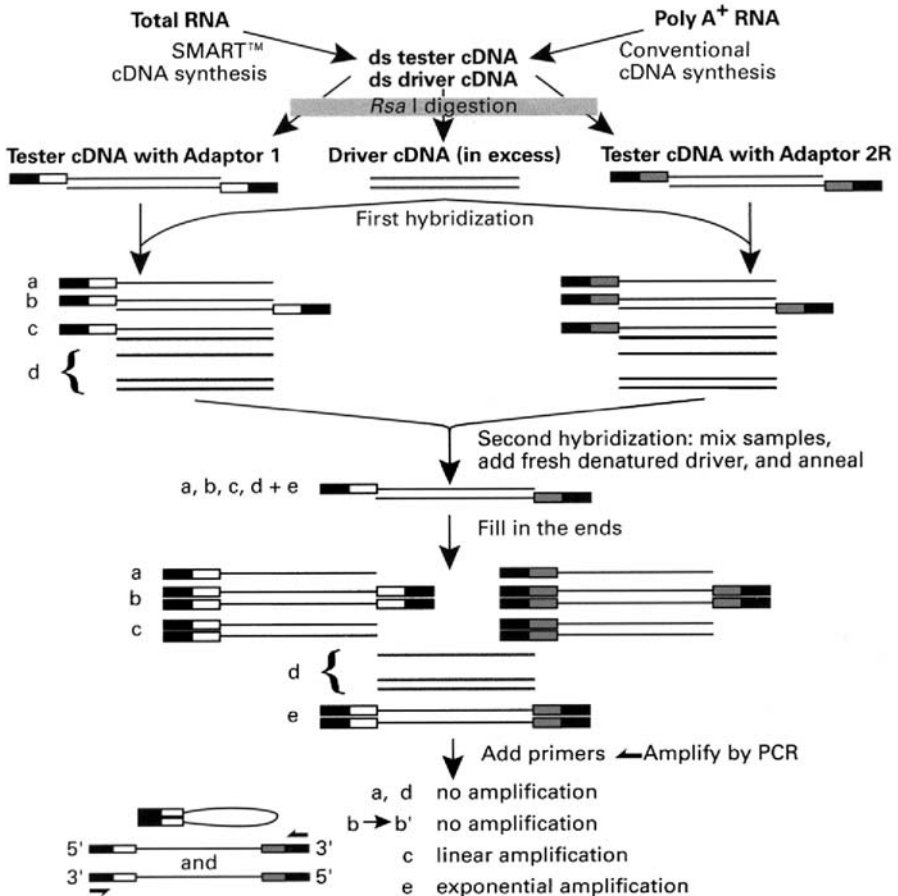


Figure 22.3 Subtraction-suppression procedure showing how the various cDNAs in the reaction, both tester and driver, are able to interact with each other. Courtesy of Clontech Laboratories (PCR-Select cDNA Subtraction system).

hours while the second hybridization may proceed overnight. Beyond 12 hours in the first hybridization the single-stranded tester molecules have a tendency to base-pair intramolecularly and with each other, making the formation of double-stranded molecules impossible in the second hybridization. The recommended hybridization periods are more than adequate for the removal of sequences shared by both nucleic acid populations.

At the end of the subtraction-suppression PCR procedures the subtracted cDNAs are ligated to plasmids, followed by transformation into *E. coli*. This effects physical separation of the individual cDNAs for a more comprehensive analysis. As plasmid preps and sequencing can be time consuming and costly,

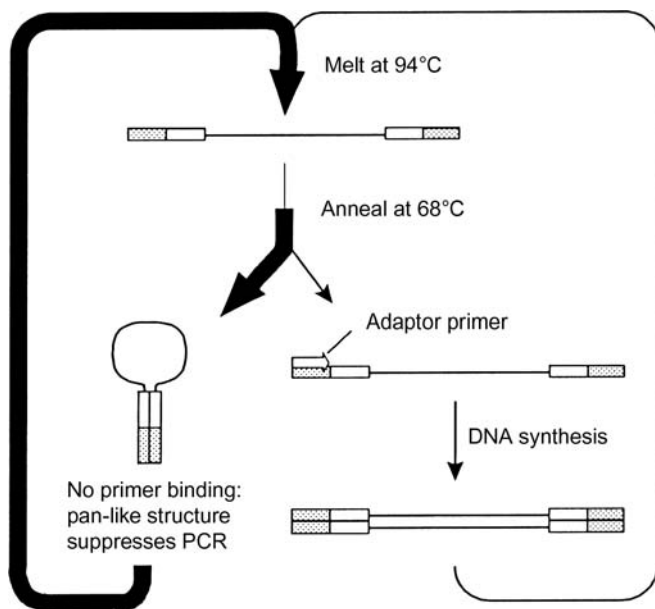


Figure 22.4 Suppression PCR. During PCR, nonspecific amplification is efficiently suppressed by the formation of a pan-like structure, while specific amplification of double-stranded cDNAs with different adaptors at both ends proceeds normally. Courtesy of Clontech Laboratories.

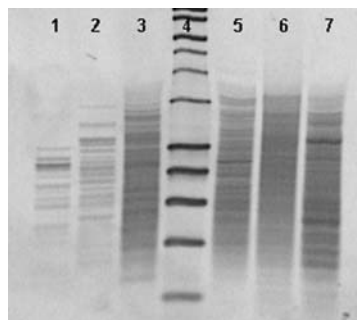


Figure 22.5 Typical results of subtraction-suppression PCR. Six RNA populations were examined for differential expression under precisely controlled environmental conditions. The samples under investigation were paired as the “A” reactions, the “B” reactions, and the “C” reactions, and a forward (A–F, etc.) and reverse (A–R, etc.) subtraction was performed for each. The subtracted products were then PCR amplified. The bands present in each lane represent differentially expressed cDNA sequences that were subsequently cloned for sequencing and further experimentation in order to confirm differential expression. Lane 1, A-forward; Lane 2, A-reverse; Lane 3, B-forward; Lane 4, molecular weight standard; Lane 5, B-reverse; Lane 6, C-forward; Lane 7, C-reverse.

a rapid, preliminary screening to confirm differential status involves colony hybridization first using an aliquot of labeled tester followed by a screening replica plated colonies with an aliquot of labeled driver. One way for efficiently managing relatively small numbers of clones (<500) involves picking the colonies with sterile toothpick and transferring them to the wells of a 96-well plate containing 200 μ l sterile LB medium supplemented with 10% glycerol and the appropriate antibiotic. After incubating the plates at 37° overnight, a replica plating tool (Fig. 22.6) is used to inoculate a fresh agar dish such that the clones grow in a geometric pattern on the agar plate, and each clone can be easily traced back to a specific well. In addition, the clones that grow after replica plating in this manner are generally quite a bit larger than the normal size of colonies that results from standard plating methods after a routine bacterial transformation. These larger colonies produce quite unambiguous signals upon hybridization to a probe. Colonies that light up with the tester probe, but not with the driver probes are assumed to be differentially expressed. At that point individual clones can be plasmid-prepped and sequenced. If the colonies were organized into glycerol-supplemented medium in 96-well plates, these plates can then be sealed with Parafilm and then stored at -80° with great confidence in the long-term viability of these clones.

Advantages and disadvantages

The suppression subtractive hybridization method overall requires a fair amount of skill and can be time-consuming relative to other methods, though the potential benefits are extraordinary. One of nice things about the system is that the subtractions can (and should) be performed in both directions, leading to the capture

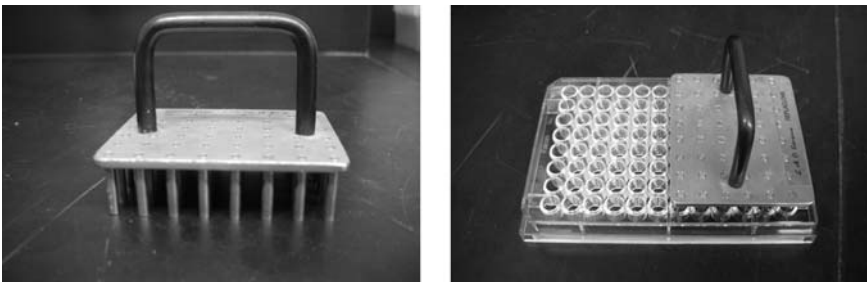


Figure 22.6 Bacterial colony replica-plating tool. The stainless steel prongs fit precisely into 48 of the 96 wells of a standard microtiter plate into which individual bacterial clones have been grown (usually 200 μ l per well). After dipping the prongs into the bacteria-containing wells, the instrument is gently placed on the surface of a standard 1.5% agar plate (with appropriate antibiotics) and the plate is then incubated as usual. The bacterial clones will grow on the plate with a very precise geometry, producing a much larger target area for hybridization. Because of the geometry and strong signals from positive clones, it will be very easy to recover positive clones from the master 96-well plate. The 96-well format is also an excellent way to archive and store clones for extended periods.

of both up- and downregulated sequences. The first- and second hybridizations very effectively normalize the enriched population of differentially expressed sequences such that one has nearly an equal chance of sequencing a cDNA from an abundant transcript or a low abundance transcript. Best of all, SSH is a truly a method for new gene discovery, since any differentially expressed sequence can be profiled, regardless of its identity. This is an advantage of microarray screening in which one assays only those sequences that were printed on the array. SSH is an excellent alternative for gene expression profiling when an array is not available.

As mentioned above, the complete SSH procedure usually takes four to five days to complete. There are many manipulations that must be performed, many of which involve generating the cDNA prior the actual subtraction itself. The normalization of sequences which occurs in the two hybridizations is efficient, but higher abundance sequences sometimes show an occasional level of redundancy greater than is observed among other sequences of lower natural abundance. Also, it is very important to confirm that the subtracted sequences are, in fact, differentially expressed, preferably using a non-PCR-based method. Overall, the system is excellent and is a valuable research tool.

Troubleshooting

1. Non-negotiables

A. High-quality RNA

- Use RNase-free technique throughout the procedures.
- Use guanidinium thiocyanate- or guanidinium HCl-based lysis buffers³.
- Perform acid-phenol extraction.
- Perform double precipitation with guanidinium lysis buffer.
- Treat samples with RNase-free DNase.
- Wash RNA pellets thoroughly and extensively with 70% ethanol prepared in nuclease-free H₂O to remove excess salt⁴.
- Properly store purified RNA.

B. Reverse transcription into cDNA and PCR

- Use an equal mass of RNA per cDNA reaction (200–300 ng).
- Recall that enzymes lacking endogenous RNase H activity may be helpful.
- Always work from a cDNA master mix.
- Always work from a PCR master mix.
- Use a hot-start PCR method, ideally in conjunction with an enzyme blend. This will ensure high processivity, specificity, fidelity, and reproducibility.

2. Common problems

A. Heavy smearing in one or more lanes:

- Optimize Mg²⁺ concentration (usually not required).
- Reduce the number of cycles.
- Decrease the amount of enzyme in the reaction.
- Slightly increase annealing temperature.

³NP-40 lysis buffers can also be used depending on the biological source, though most protocols recommend the use of a chaotropic guanidinium lysis reagent.

⁴Cationic salts bind dNTPs quantitatively and can cause havoc in the cDNA synthesis and PCR reactions.

B. No product:

- Check the integrity of the starting RNA!
- Add an RNase inhibitor (e.g., RNasin™) to the cDNA synthesis reaction.
- Test the components of the RT and PCR reactions, especially the enzymes.
- Run *all* recommended control reactions.
- Check off each component as the reactions are pipetted together.
- Check PCR cycle parameters.

C. Low yield:

- Incomplete homogenization or lysis of samples could be the cause.
- Ensure that the final RNA pellet is incompletely redissolved.
- If $A_{260}/_{280} < 1.65$, the sample is probably contaminated with protein.
- Check for RNA degradation:
 - Tissue was not immediately processed/frozen after removing from the animal. Consider the use of RNAlater.
 - Purified RNA was stored at -20° instead of -70° .
 - Aqueous solutions or tubes were not RNase-free.

D. DNA contamination:

- Recall that this may occur if the sample was not DNase-treated (should be a standard procedure).

E. Non-reproducibility of data:

- Use thin-walled tubes (200 μ l work well); for thick walled tubes, extend cycle parameters.
- Consider hot start PCR (described in Chapter 18).
- Prepare single, large-volume master mix per cell type and aliquot for each reaction.
- Match T_m of upstream and downstream primer pairs as closely as possible.
- Consider enzyme mixtures designed for long-range PCR.
- Consider elimination of mineral oil: use a thermal cycler with a heated lid.
- Optimize and run reactions in the same thermal cycler (not just the same model number).
- Non-reproducibility can occur when RNA is not isolated by the same method or on the same day. For example, if the cells are in different parts of the cell cycle or the tissues are isolated from biochemically different individuals, the data may show considerable variation.

3. General suggestions

- Keep denaturation steps as short as possible.
- Keep denaturation temperature as low as possible (90 – 92°).
- Elongation temperature can be reduced to 68° .
- Extend elongation time for each successive cycle (5 – 20 s per cycle).
- Accurate pipetting is critical for reproducibility. Micropipettors should be recalibrated at the onset of the study, and at regular intervals thereafter. Use barrier tips to avoid PCR carryover contamination.
- In some cases, very low abundance mRNA can be superinduced by adding a sub-cytotoxic concentration of cycloheximide (10 ng/ml) to the cell culture growth medium and incubating for 4 to 6 hours before cell lysis and RNA isolation. When the protein translation apparatus is inhibited in this manner, mature mRNAs tend to accumulate in the cytoplasm. However, some cells, especially cultures of normal diploid cells, will not tolerate this treatment.

- Unlike many PCR-associated techniques, transcript subtraction approaches seem to be quite sensitive to the quality of the starting material. The highest quality RNA is generally prepared in guanidinium-containing buffers.
- Silanized microcentrifuge tubes may be helpful in minimizing losses during the various manipulations because both cDNA and mRNA adhere to the inside of microfuge tubes. Silanizing is especially helpful when only very small quantities of irreplaceable samples are available (see Appendix I for protocol).

Non-subtractive methods

An alternative strategy to profiling gene expression in two or more samples involves coupling standard methods for cDNA synthesis to PCR-based amplification using large combinations of short primers. In so doing, transcripts common to the RNA populations under investigation are manifested as PCR products of identical size when compared side by side using polyacrylamide gel electrophoresis (PAGE) when identical primer sets are used. More importantly, this method facilitates the identification of bands in one lane that are *absent* or of *different abundance* in the corresponding location of the adjacent lane(s). This strategy for gene expression profiling is more about selective amplification rather than subtraction.

*mRNA Differential display*⁵

Of the several variations on this standard method, the most scientifically correct terminology to describe this technique is mRNA differential display (Liang and Pardee, 1992; Liang *et al.*, 1992; Liang *et al.*, 1993; Mou *et al.*, 1994; Liang *et al.*, 1995; Zhao *et al.*, 1995), though it is also referred to as differential display PCR, DDPCR⁶, or the more cumbersome DDRTPCR (Bauer *et al.*, 1993). Other methods that have evolved from this general approach, but with a subtle variations in primer design, are the amplified restriction fragment length polymorphism (AFLP) technique (Vos *et al.*, 1995), and RNA arbitrarily primed PCR (RAP-PCR) (Welsh *et al.*, 1992; McClelland and Welsh, 1994). Although subtly different from one another, each of these techniques essentially accomplishes the same end albeit with varying levels of coverage with respect to transcript characterization.

Background

The strategy behind DDPCR is quite simple: incorporation of traditional synthesis of cDNA from purified RNA, followed by amplification of the products

⁵The mRNA differential display process is governed by exclusive patent rights granted to GenHunter Corporation (Nashville, TN). As such, GenHunter is the only authorized vendor of kits for performing differential display PCR. Visit www.genhunter.com for more information.

⁶When using this shorthand, be certain to use uppercase letters. The shorthand ddPCR strongly suggests that the writer is referring to dideoxy (dd-) sequencing by PCR.

of the first-strand cDNA reaction. Unlike traditional cDNA synthesis, several cDNA reactions are performed for each sample, each of which is supported by a unique combination of downstream primers and upstream primers, which, in any given reaction tube, will support the amplification of only a subset of all possible PCR products. Upon PAGE analysis, DDPCR products manifest themselves as being highly abundant, not at all abundant, and everywhere in between. The data make the greatest sense when all of the PCR products generated using the same downstream primer are electrophoresed on a single gel. "State A" PCR reactions are loaded adjacent to "State B" PCR reactions, in order of upstream primer used, an example of which is shown in Table 22.1. This permits the direct side-by-side comparison of the products from each pair of reactions. Products of identical molecular weight will be observed when cells in state A and cells in state B are producing the same transcript (Fig. 22.7; the appearance of band without a corresponding band in the adjacent lane suggests differential expression (Fig. 22.8) Conveniently, 48-tooth combs used for DNA sequencing permit the loading of multiple primer combinations, using a single downstream primer, on the same gel (e.g., 2 cell samples \times 1 downstream primer \times 24 upstream primers). The use of a molecular weight marker is not required on this gel, the purpose of which is merely to highlight the regulated sequences. Molecular weight standards are more valuable following band

Table 22.1 Gel Loading Strategy for Differential Display PCR

Lane	cDNA source	Downstream primer	Upstream primer
1	State A	3	1
2	State B	3	1
3	State A	3	2
4	State B	3	2
5	State A	3	3
6	State B	3	3
7	State A	3	4
8	State B	3	4
9	State A	3	5
10	State B	3	5

11–48 *Same strategy*

PCR products generated using identical primer combinations are loaded into adjacent lanes, permitting the direct side-by-side comparison of the variety of cDNAs amplified. Bands of equal molecular weight in adjacent lanes correspond to sequences transcribed in both cell types. Differences in intensity, including the complete absence of a band in one of the lanes, suggest differential modulation of specific genes. Standard-sized DNA sequencing gels favor excellent resolution and a 48-tooth comb permits loading several upstream- and downstream primer combinations for both cell types.

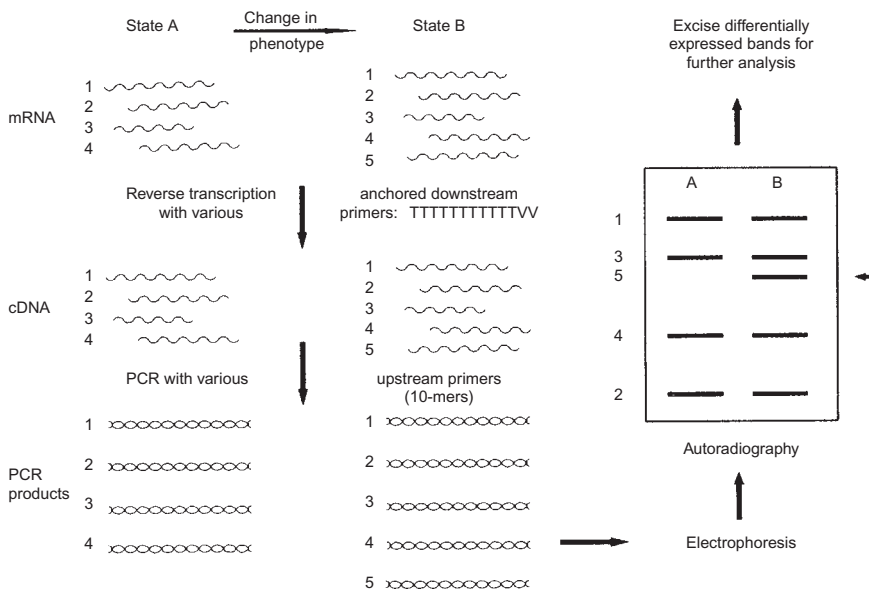


Figure 22.7 Schematic showing the general method used for mRNA differential display. Transcripts produced in both cell types are displayed simultaneously as PCR products when identical primer combinations are used. The presence of a PCR product from only one sample strongly suggests differential transcriptional activity of the corresponding gene.

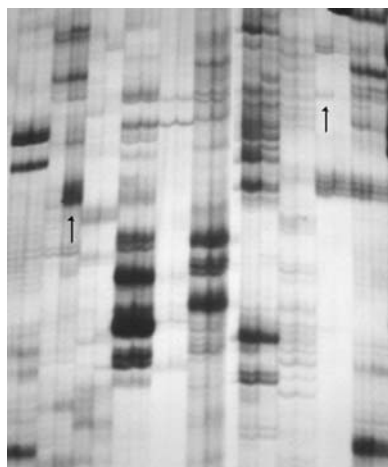


Figure 22.8 mRNA differential display. Typical appearance of differential display PCR products following electrophoresis and autoradiography of the first round amplification. Identical bands appearing in adjacent lanes represent genes transcribed in common. Bands unique to one lane in a set are potentially differentially expressed sequences and are, therefore, candidates for further characterization. Notice the differing intensities of the various product pairs, directly reflecting the relative abundance of the corresponding transcripts.

retrieval and subsequent characterization. Taken together, these steps constitute a structured method for simultaneous examination of all transcribed sequences.

Virtually any model in which experimental manipulation is likely to cause changes in the cellular biochemistry is a suitable candidate for evaluation using DDPCR technology. mRNA differential display affords a side-by-side comparison of mRNA from different sources and is excellent for discerning both the induction and repression of genes. Examples include, but certainly are not limited to, the study of gene modulation (1) in the presence of pharmacological concentrations of growth factors or drugs; (2) when comparing quiescent to senescent cells; (3) as a function of seasonal changes in fruits and vegetables; (4) in response to the expression of transfected foreign genes; (5) as a function of differentiation; (6) for the isolation of new members of multigene families; and (7) for the discernment of transcriptionally regulated sequences versus those that are posttranscriptionally regulated (relates to method of RNA preparation – refer to Fig. 1.8 in Chapter 1).

Protocols

The major steps in DDPCR are as follows:

1. **RNA isolation.** Identification of modulated sequences by DDPCR begins with the isolation of extremely high-quality RNA. In particular, the isolation of RNA by guanidinium-acid-phenol is one of the better, more widely accepted, purification methods for analysis by DDPCR. Investigators may wish to explore the protocols found in Chapters 2–4, or use one of the numerous commercial guanidinium-containing reagents to accomplish this isolation. Of profound importance is the removal of all traces of contaminating genomic DNA by DNase-treating the sample. Remember that primers will base-pair to any complementary molecule, DNA or RNA. The efficient removal of DNA is validated by the no-reverse transcriptase control reaction, described later.
2. **cDNA downstream primer design.** The standard approach to cDNA synthesis intended to support analysis by DDPCR is to perform several cDNA synthesis reactions for each unique RNA sample, reverse transcribing only a fraction of all possible messages into cDNA in any one reaction tube. Thus, each reverse transcriptase reaction is *intended* to generate subpopulations of cDNA so as to resolve discrete bands by electrophoresis, rather than a glorious smear. The key to accomplishing this lies with the use of anchored downstream primers which often have the generic structure (Liang *et al.*, 1994):

$$5' \text{ TTTTTTTTTTV}$$

or

$$5' \text{ TTTTTTTTTTV}$$

where V = A, C, or G.

The 3' nucleotide degeneracy ensures the reverse transcription of all possible mRNAs; further, the 3' nucleotide anchor ensures that the primer will base pair at the 5'-most end of the poly(A)⁺ tail: an oligo(dT) primer (5' TTTTTTTTTTTTTT)

can base-pair well into the body of the poly(A) tail, hundreds of nucleotides away from the coding portion of the transcript. Some investigators prefer downstream primers with 3' dinucleotide degeneracy, and this is a matter of personal preference.

3. **cDNA synthesis.** The first manipulation of RNA destined for differential display is the synthesis of first-strand cDNA. First-strand cDNA is traditionally performed in the presence of ^{35}S or ^{32}P , though there are non-radioactive variants of the DDPCR technique as well (Lohman *et al.*, 1995). The first-strand cDNA should be synthesized and store appropriately so that long-term mRNA stability of potentially irreplaceable samples does not become an issue. It is very important to synthesize the longest cDNAs possible: if the efficiency of the first strand synthesis is compromised, then shorter first-strand molecules may not be long enough to base-pair with the upstream primers and will thus not be amplified by PCR. Because of the intrinsically labile nature of RNA, the small mass of RNA used for each cDNA synthesis reaction, and the peculiarities of this technique, strong consideration should be given to the use of an RNase H⁻ reverse transcriptase (see Chapter 17 for details).
4. **PCR amplification.** The design of the upstream primers is generally less structured than that of the downstream primers. The upstream primers are typically 10-mers or 13-mers and should be thought of as what this Author likes to refer to as "semi-non-specific". Given the relatively relaxed primer annealing conditions, and the fact that the primers are rather short, only seven or eight of the bases in any given upstream primer are expected to anneal. While this relatively relaxed stringency may at first be unsettling, it is a necessary component of the DDPCR approach: the idea is to be able to amplify *any* message that nature can generate without *a priori* knowledge of its sequence. Ergo, it is not necessary to worry about all possible 10-mer (4^{10}) or 13-mer (4^{13}) combinations. Precisely because of the nature of the primers, hot-start PCR may be helpful.

The upstream primers support the subsequent synthesis of the second strand cDNA and, in conjunction with the downstream primer, will support the amplification of the cDNA by PCR. It is clear from the preceding information that there is potentially a large number of PCR reactions that can be performed. It would not be reasonable to consider running all possible primer combinations simultaneously. Instead, one may store some of the first-strand cDNA synthesis reactions at -80° . Should interesting PCR products become manifest by DDPCR with the first set of samples, the bands can be fully characterized, and the remaining PCR reactions performed at a later date.

The importance of suitable controls cannot be overstated because no molecular biology technique is perfect. First, it is incumbent upon the investigator to always run a no-reverse transcriptase (RT-) control whenever RNA is reverse transcribed. Failure to generate product will support the notion that the RNA prep was free of contaminating, competing genomic DNA template material. If PCR product is generated in an RT- control reaction, the investigator must DNase-treat the sample (with RNase-free DNase I) before moving forward. One should assume that all RNA samples are tainted with genomic DNA and it is a tremendous time saver to automatically DNase-treat all RNA samples during the isolation protocol rather than having to back-track.

Second, be certain to run control reactions with only one primer, to show that the primers are not self-priming. One method that is useful in this regard is to kinase (5' end label) the downstream primers. Thus, the only PCR products visible by

autoradiography are those generated with an upstream *and* downstream primer, as opposed to upstream primers alone.

Third, a simple way to ensure authenticity of a differentially amplified product is to assay two concentrations of the same RNA sample. A true positive, one which will generate a signal with a least two starting concentrations of the same RNA can be further investigated. This precludes wasting time investigating artifactual variations, that is, differences in signal intensity due to the way the samples were pipetted together or the way the gel was loaded.

5. **Pilot study.** In view of the potential for technical difficulties, it is strongly recommended that a pilot study, or at least a pilot set of reactions, be developed before actually investing labor and materials on samples that may well be irreplaceable, especially if reproducibility is an issue. For example, it can be helpful to electrophorese a small aliquot of the completed reactions through 3% agarose and then stain it with SYBR Green or ethidium bromide (Fig. 22.9) before making the effort to run and analyze the DDPCR products on a polyacrylamide gel. While it is true that these dyes and a 3% gel offer very limited sensitivity and resolving power, the objective here is to simply demonstrate that the reactions actually worked before moving forward.

Typical DDPCR reactions generate usable products that range from 75 to 800 bp, while products larger than 800 bp begin to show signs of compression. Clearer definition may be given to the larger molecular weight products by double loading the gel, as is often done for DNA sequencing purposes. PCR products are visualized by autoradiography. It is usually a good idea to identify internal markers within a given set of reactions electrophoresed on the same gel; that is, look for the superabundant sequences in the same relative position on the gel each time the experiment is performed. A very short exposure to X-ray film will show these strong signals only, and not the lesser-strength signals: this is a good way to make a fingerprint of the reaction and is way to check efficiency each time a set of samples is assayed

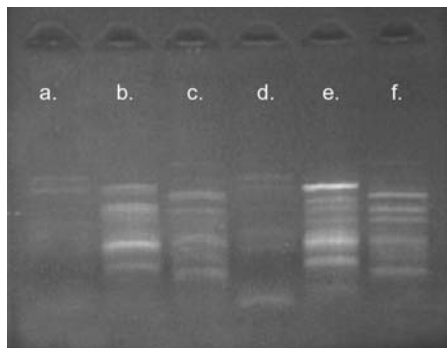


Figure 22.9 Appearance of DDPCR products after first round amplification when stained with ethidium bromide. Aliquots of six samples were electrophoresed in a 3% agarose gel supplemented with 1 $\mu\text{g}/\text{ml}$ ethidium bromide. This diagnostic can be used as a preliminary indicator of the efficiency of the reactions. Notice that, even with this very insensitive method of visualization, some differences are obvious among the samples.

because these abundant “landmark” sequences should show up each time an experiment is repeated. Differential display PCR is at least semiquantitative⁷ at this level, reflecting the complexity of the RNA starting material: signal intensities of electrophoretically resolved products are in proportion to the abundance of the original mRNA template material. Finally, bands of interest are cut out of the acrylamide gel using the autoradiogram as a guide. Gel slices are heated in sterile H₂O to liberate the PCR product(s) after which it is reamplified using the same primer combination that supported its synthesis in the DDPCR reactions. These products are then cloned into a suitable vector for sequencing (Wang and Feuerstein, 1995; Reeves *et al.*, 1995), confirmation of differential expression (Zhang *et al.*, 1996) and any of a variety of downstream applications.

6. **Confirmation of differential expression.** After DDPCR is run once using all primer combinations, a second run is conducted using only those primer pairs that show differences in the first run. To an extent, this approach favors the elimination of false positives. Of course, the exact experimental milieu that defined the first run should be duplicated as closely as possible, to afford an opportunity to reproduce the data. One should not expect, however, that any two runs will be precisely reproducible because (1) cell culture conditions are impossible to duplicate precisely; and (2) tissue samples from different patients are exactly that – biochemically different organisms. While many similarities are to be expected, many differences should not be unexpected. To maximize the reproducibility of different runs from the same sample, the investigator should aliquot the RNA during the initial isolation procedure such that the RNA is precipitated in several aliquots, only one of which is centrifuged and used as cDNA template material. Upon completion of the entire DDPCR protocol, the investigator will then have access to a fresh aliquot of RNA that was isolated on the same day, by the same person, and that represents precisely the same biochemical status of the cells.

At present, there are four reliable methods for assessing whether a sequence purported to be differentially expressed, using the differential display method, truly represents a biochemical change in cell: Northern analysis (Chapter 11), ribonuclease protection assay (RPA; Chapter 15), nuclear runoff assay (Chapter 16), or real-time PCR. If possible, the confirmation is all the more convincing when demonstrated using a non-PCR method.

The isolation of differentially modulated sequences opens many possibilities for gaining a clearer understanding of this aspect of gene regulation in the cell. Isolated DNA sequences can be reamplified repeatedly, constituting a limitless source of nucleic acid probe; PCR with merely 15 cycles, in the presence of [α -³²P] dNTP, is all that is required. Probe sequences can be used to perform gene expression assays using previously untested RNA samples, or to screen libraries to identify larger portions of the gene as well as flanking regulatory elements. Well-characterized sequences may also be cloned into expression vectors in order to observed the effects of over-production of a particular protein.

⁷There is a limit to the sensitivity of DDPCR, which might be described as semiquantitative at best. Subsequent analyses, such as nuclease protection (Chapter 15) or real-time PCR (Chapter 19) are used to measure changes in gene expression in a much more quantitative manner.

The gene sequence can also be mutated (site-directed mutagenesis) to determine the detriment to the cell when a misfolded or truncated protein is produced. The possibilities are endless and mind-boggling.

Advantages and disadvantages

There are clear advantages and disadvantages associated with the DDPCR method which have been summarized in Table 22.2. Succinctly, the requirement for small input RNA mass is nice, though it is strongly recommended that a larger mass of RNA be purified and stored as an ethanol precipitate. This ensures that precisely the same RNA template, from the same biochemical source, isolated on the same day by the same investigator, will be available when the time comes to repeat the experiment and, hopefully, reproduce the data. Moreover, the selection of poly(A)⁺ RNA is neither required nor recommended, which may actually serve to further under-represent low abundance transcripts.

Inasmuch as phenotypic change unquestionably involves several genes, another key advantage of DDPCR is that the mechanics of the assay favor the simultaneous display of all transcripts present in the cell at the moment of cell

Table 22.2 Advantages and Disadvantages of Differential Display PCR

Advantages

- The reaction is simple, fast, sensitive method for detecting specific transcripts.
- The method is a versatile and semiquantitative method.
- It is a systematic approach for the identification of differential gene expression.
- Poly(A)⁺ selection is neither required nor recommended.
- There is a simultaneous display of all up- and downregulated genes.
- Microgram quantities of RNA are required (often less than 3 µg).
- Data are reproducible, though this is user-dependent.
- Tool for new gene discovery.

Disadvantages

- PCR products need to be sequenced for identification (compare to microarrays).
 - Very pure samples of RNA are required.
 - The reaction is very sensitive to contaminants, especially genomic DNA.
 - The location of known sequences on the resulting gel cannot be predicted.
 - Large numbers of PCR reactions are necessary to fully compare a pair of samples.
 - High incidence of false positives (20–30%) because of fairly low-stringency annealing and other variables.
 - False positives are also observed after reamplification.
 - Purified cDNAs are often contaminated with unrelated cDNA sequences.
 - The traditional format favors amplification of sequences corresponding to the 3' end of transcripts.
-

lysis. The speed and sensitivity of PCR facilitates the detection of very-low-abundance transcripts in a relatively short time frame. A most important concept here is that differences in gene expression are manifested in the form of PCR products, though the investigator has no idea what these PCR products are until they have been sequenced.

RNA used for differential display is often prepared by guanidinium-acid-phenol extraction, which is a perfectly legitimate method for RNA recovery from a variety of biological sources. Recall, however, that such chaotropic reagents yield mRNA from the cytoplasm mixed with unspliced hnRNA upon cellular disruption. Thus, there exists the potential that a sequenced PCR product may show both exon and intron sequences. In addition, it is often the case that two different size cDNAs, corresponding to the same transcript, are identified on the same gel. This occurs because in both cases all of the molecules of a specific transcript are reverse transcribed by the same downstream primer, though two or more upstream primers (different reaction tubes) find complementary sites on the cDNA, thereby supporting amplification by PCR. Although this seems to frustrate many, the phenomenon should be viewed as a means of maximizing the chances that any given transcript will be amplified.

Troubleshooting

Among the more common complaints engendered by the use of DDPCR is the seemingly non-reproducible nature of the reactions, and there appear to be limitless theories promulgated to explain this. As with all PCR-based methods, the assay is exquisitely sensitive to the precise concentration of all reaction components; thus, accurate pipetting and the use of master mixes for both cDNA synthesis and PCR are essential. The DDPCR approach works nicely when performed properly, and the following suggestions outline what may be done to minimize aggravation and maximize reproducibility and productivity.

1. Non-negotiables

A. High-quality RNA

- Use RNase-free technique throughout the procedures.
- Use guanidinium thiocyanate- or guanidinium HCl-based lysis buffers⁸.
- Perform acid-phenol extraction.
- Perform double precipitation with guanidinium lysis buffer.
- Treat samples with RNase-free DNase I.
- Wash RNA pellets thoroughly and extensively with 70% ethanol prepared in nuclease-free H₂O to remove excess salt⁹.
- Properly store purified RNA.

B. Reverse transcription into cDNA and PCR

- Use equal mass of RNA per cDNA reaction (200–300 ng).

⁸NP-40 lysis buffers can also be used depending on the biological source, though most protocols recommend the use of a chaotropic guanidinium lysis reagent.

⁹Cationic salts bind dNTPs quantitatively and can cause havoc in downstream reactions.

- Recall that enzymes lacking endogenous RNase H activity may be helpful.
- Always work from a cDNA master mix.
- Always work from a PCR master mix.

2. Common problems

A. Heavy smearing in one or more lanes:

- Optimize Mg^{2+} concentration; decrease or increase incrementally by 0.1 mM.
- Reduce the number of cycles to as few as 30.
- Decrease the amount of enzyme in the reaction.
- Slightly increase annealing temperature.

Note: In some cases, smearing may be reduced if Cycles 1–5 are annealed at 40° and Cycles 6–40 are annealed at 45°. Be aware, however, that this strategy may reduce the variety of bands by as much as 25–30%.

B. No product:

- Check the RNA!
- Add an RNase inhibitor (e.g., RNasin™).
- Test the components of the RT and PCR reactions, especially the enzymes.
- Check off each component as the reactions are pipetted together.
- Check PCR cycle parameters.

C. Low yield:

- Consider that incomplete homogenization or lysis of samples may have occurred.
- Check to see if final RNA pellet incompletely redissolved.
- Determine if $A_{260}/_{280} > 1.65$. A lower ratio suggests protein or other contaminants.
- Check for RNA degradation:
 - Tissue was not immediately processed/frozen after removing from the animal.
 - Purified RNA was stored at -20° instead of -80° .
 - Aqueous solutions or tubes were not RNase-free.

D. DNA contamination:

- This occurs if the sample was not DNase-treated.

E. Non-reproducibility of data:

- Use thin-walled tubes (200 μ l work well); for thick walled tubes, extend cycle parameters.
- Consider hot start PCR (described in Chapter 18).
- Prepare single, large-volume master mix per cell type and aliquot for each reaction and be attentive to master mix composition.
- Verify primer combinations used with each sample.
- Match T_m of upstream and downstream primer pairs as closely as possible.
- Consider enzyme blends designed for long PCR.
- Consider elimination of mineral oil: use a thermal cycler with a heated lid.
- Optimize reactions in the same thermal cycler (not same model number).
- Non-reproducibility can occur when RNA is not isolated by the same method or on the same day. For example, if the cells are in different parts of the cell cycle or the tissues are isolated from biochemically different individuals, the data may show considerable variation.

3. General suggestions

- Keep denaturation steps as short as possible.
- Keep the denaturation temperature as low as possible (90–92°).
- The elongation temperature can be reduced to 68°.
- The elongation time can be extended for each successive cycle (5–20 s per cycle).

- Accurate pipetting is critical for reproducibility. Have micropipettors recalibrated before beginning and at regular intervals thereafter. Use barrier tips to prevent PCR carryover contamination due to aerosol formation.
- The mRNA differential display technique is especially sensitive to the exact input mass of RNA. As with most RT-PCR reactions “less is more”, meaning that it is counter-productive to overload the reaction with starting material (200–300 ng of RNA is optimal).
- Unlike many PCR-associated techniques, differential display seems to be quite sensitive to the quality of the starting material. The highest quality RNA is generally prepared in guanidinium-containing buffers.
- Silanized microcentrifuge tubes may be helpful in minimizing losses during the various manipulations because cDNA and PCR products are present in picogram quantities (or less) prior to reamplification (see Appendix I for protocol).

References

- Bassett, C.L., M.E. Wisniewski, T.S. Arlip, J.L. Norelli, J. Renaut, and R.E. Farrell, Jr. (2006). Global analysis of genes regulated by low temperature and photoperiod in peach bark. *J. Am. Soc. Hort. Sci.* **131**, 551–563.
- Bauer, D., H. Muller, J. Reich, H. Riedel, V. Ahrenkiel, P. Warthoe, and M. Strauss. (1993). Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Res.* **21**, 4272–4280.
- Bernardi, G. (1971). Chromatography of nucleic acids on hydroxyapatite columns. *Methods Enzymol.* **21**, 95–139.
- Coche, T., M. Dewez, and M.C. Beckers. (1994). Generation of an unlimited supply of subtracted probe using magnetic beads and PCR. *Nucleic Acids Res.* **22**, 1322–1323.
- Cochran, B.H., P. Zumstein, J. Zullo, J. Rollins, M. Mercola, and C.D. Stiles. (1987). Differential colony hybridization: molecular cloning from a zero data base. *Methods Enzymol.* **147**, 64–85.
- Fornace, A.J. and J.B. Mitchell. (1986). Induction of B2 RNA polymerase III transcription by heat shock: enrichment for heat shock induced sequences in rodent cells by hybridization subtraction. *Nucleic Acids Res.* **14**(14), 5793–5811.
- Hedrick, S.M., D.I. Cohen, E.A. Nielsen, and M.M. Davis. (1984). Isolation of cDNA clones encoding T-cell specific membrane-associated proteins. *Nature* **308**, 149–153.
- Houge, G. (1993). Simplified construction of a subtracted cDNA library using asymmetric PCR. *PCR Methods and Applications* **2**(3), 204–209.
- Kulesh, D.A., D.R. Clive, D.S. Zarlenga, and J.J. Greene. (1987). Identification of interferon-modulated proliferation-related cDNA sequences. *Proc. Natl. Acad. Sci.* **84**, 8453–8457.
- Lambert, K.N. and V.M. Williamson. (1993). DNA library construction from small amounts of RNA using paramagnetic beads and PCR. *Nucleic Acids Res.* **21**, 775–776.
- Liang, P. and A.B. Pardee. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.
- Liang, P., L. Averboukh, K. Keyomarsi, R. Sager, and A.B. Pardee. (1992). Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer Res.* **52**, 6966–6968.

- Liang, P., L. Averboukh, and A.B. Pardee. (1993). Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Res.* **21**, 3269–3275.
- Liang, P., W. Zhu, X. Zhang, Z. Guo, R.P. O'Connell, L. Averboukh, F. Wang, and A.B. Pardee. (1994). Differential display using one-base anchored oligo-dT primers. *Nucleic Acids Res.* **22**, 5763–5764.
- Liang, P., D. Bauer, L. Averboukh, P. Warthoe, M. Rohrwild, H. Muller, M. Strauss, and A.B. Pardee. (1995). Analysis of altered gene expression by differential display. *Methods Enzymol.* **254**, 304–321.
- Lisitsyn, N., N. Lisitsyn, and M. Wigler. (1993). Cloning the differences between two complex genomes. *Science* **259**, 946–951.
- Lohman, J., H. Schickle, and T. Bosch. (1995). REN Display, a rapid and efficient method for nonradioactive differential display and mRNA isolation. *BioTechniques* **18**, 200–202.
- López-Fernández L.A. and J. del Mazo. (1993). Construction of subtractive cDNA libraries from limited amounts of mRNA and multiple cycles of subtraction. *BioTechniques* **15**(4), 654–658.
- McClelland, M. and J. Welsh. (1994). RNA fingerprinting by arbitrarily primed PCR. *PCR Methods Appl.* **4**, S66–S81.
- Mou, L., H. Miller, J. Li, E. Wang, and L. Chalifour. (1994). Improvements to the differential display method for gene analysis. *Biophys. Res. Commun.* **199**, 564–569.
- Norelli, J.L., R.E. Farrell, Jr., C.L. Bassett, A.M. Baldo, D.A. Lalli, H.S. Aldwinckle, and M.E. Wisniewski. (2009). Rapid transcriptional response of apple to fire blight disease revealed by cDNA suppression subtractive hybridization analysis. *Tree Genet. Genomes* **5**, 27–40.
- Orr, S.L., E. Gese, and L. Hood. (1992). A new approach to understanding T cell development: the isolation and characterization of immature CD4⁺, CD8⁻, CD3⁻ T cell cDNAs by subtraction cloning. *Mol. Biol. Cell* **3**, 761–773.
- Reeves, S.A., M. Rubio, and D.N. Louis. (1995). General method for PCR amplification and direct sequencing of mRNA differential display products. *BioTechniques* **18**(1), 18–20.
- Rodriguez, I.R. and G.J. Chader. (1992). A novel method for the isolation of tissue-specific genes. *Nucleic Acids Res.* **20**, 3528.
- Sargent, T.D. (1987). Isolation of differentially expressed genes. *Methods Enzymol.* **152**, 423–432.
- Sargent, T.D. and I.B. Dawid. (1983). Differential gene expression in the gastrula of *Xenopus laevis*. *Science* **222**, 135–139.
- Sharma, P., A. Lönneborg, and P. Stougaard. (1993). PCR-based construction of subtractive cDNA library using magnetic beads. *BioTechniques* **15**(4), 610–611.
- Siebert, P.D., A. Chenchik, D.E. Kellogg, K.A. Lukyanov, and S.A. Lukyanov. (1995). An improved method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**, 1087–1088.
- Thomas, M.G., S.A. Hesse, S. Al-Mahdawi, T.D. Bui, C.F. Moniz, and F. Farzaneh. (1994). Procedure of second-round differential screening of cDNA libraries. *BioTechniques* **16**(2), 229–232.
- Travis, G.H. and J.G. Sutcliffe. (1988). Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low-abundance monkey cortex-specific mRNAs. *Proc. Natl. Acad. Sci.* **85**, 1696–1700.

- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van De Lee, and M. Hornes. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**, 4407–4414.
- Wang, X. and G.Z. Feuerstein. (1995). Direct sequencing of DNA isolated from mRNA by differential display. *BioTechniques* **18**(3), 448–452.
- Welcher, A.A., A.R. Torres, and D.C. Ward. (1986). Selective enrichment of specific DNA, cDNA, and RNA sequences using biotinylated probes, avidin, and copper-chelate agarose. *Nucleic Acids Res.* **14**, 10027–10044.
- Welsh, J., K. Chada, S.S. Dalal, R. Cheng, D. Ralph, and M. McClelland. (1992). Arbitrarily primed PCR fingerprinting of RNA. *Nucleic Acids Res.* **20**, 4965–4970.
- Zhang, H., R. Zhang, and P. Liang. (1996). Differential screening of gene expression difference enriched by differential display. *Nucleic Acids Res.* **24**, 2454–2456.
- Zhang, M., Z. Zou, N. Maass, and R. Sager. (1995). Differential expression of elafin in human normal mammary epithelial cells and carcinomas is regulated at the transcriptional level. *Cancer Res.* **55**, 2537–2541.
- Zhao, S., S.L. Ooi, and A.B. Pardee. (1995). New primer strategy improves precision of differential display PCR. *BioTechniques* **18**(5), 842–850.

23 RNAi: Take a RISC – Role the Dicer

Rationale	539	RNAi and alternative transcript splicing	552
Essential RNAi nomenclature	541	In vitro and in vivo issues	553
RNA interference – how it works	542	RNAi validation	555
siRNA approach	545	RT-PCR approaches	556
shRNA approach	546	Northern analysis	556
siRNA delivery methods into mammalian cells	548	Western analysis	556
miRNA	549	Applications	557
Effective design of siRNAs	550	References	560

Rationale

In the mid-1980s, a series of papers were published describing a novel method for *in vitro* synthesis of DNA (Saiki *et al.*, 1985; Mullis *et al.*, 1986; Mullis and Faloona, 1987). The method was a form of primer extension, a fairly well-known mapping and probe-labeling technique at the time. The innovation associated with the method was the repetitive application of a series of heating and cooling steps, resulting in impressive levels of DNA synthesis. This method, now commonly known as the polymerase chain reaction or simply PCR, revolutionized the fledgling biotechnology industry. PCR changed everything and has had a positive impact on literally all facets of research in life science, and far beyond. Forensics, basic and applied research, diagnosis of infectious disease, and maternity and paternity testing all owe their current level of accuracy, sensitivity, and sophistication to PCR. Now, a remarkable new tool has emerged that is able to suppress expression of specific endogenous genes through the use of double-stranded RNA (dsRNA). In the few short years since the demonstration of the power of this technique in *Caenorhabditis elegans* (Fire *et al.*, 1998) and subsequently in human- and other mammalian cells (Hammond *et al.*, 2000; Elbashir *et al.*, 2001a) and in plants, refinements in the methodology are impacting research much as PCR did 20 years ago. This new and emerging technology is

called *RNA interference*¹ (RNAi). It is also referred to as gene knockdown, post-transcriptional gene silencing (in plants), cosuppression (older term, in plants), and quelling (in fungi).

RNAi is a natural phenomenon that was considered an oddity when it was first observed in petunias (Napoli *et al.*, 1990; van der Krol *et al.*, 1990). It is now known to exist in many organisms as an endogenous means of protecting against viruses and transposons, molecular invaders that would otherwise plague a host genome and wreak havoc. This method of protecting the integrity of the eukaryotic genome is highly evolutionarily conserved: dsRNA entering the cell is targeted for immediate destruction. Cellular processes mediated by RNAi include the natural turnover of both wild-type and mutant mRNAs, translational regulation during the development of an organism, and undoubtedly other regulatory mechanisms in the cell that have not yet come to light. There is also evidence to suggest that the RNAi process may be related to DNA methylation (Wessenecker *et al.*, 1994; Pélissier and Wessenecker, 2000; Merrett *et al.*, 2000; Llave *et al.*, 2000) and, astonishingly, may have a role in the upregulation of gene expression (Li *et al.*, 2006; Jankowski *et al.*, 2007).

The use of RNAi as a means of studying the effects of gene expression in a cell or in an organism is occasionally called reverse genetics, the goal of which is to determine the consequences for a cell when a protein is *not* produced. In other words, the function of a gene can often be discovered by silencing it. This has profound ramifications in the realm of developmental biology, the study of the progression of the disease state, and the treatment of infectious- and other diseases that result from inappropriate protein expression, for example, through gain-of-function mutations. The major strength of RNAi is that it permits the study of the function of one gene at a time and, if properly designed, allows one to do so over an extended period.

RNAi has become a mainstream tool for both basic- and applied research, particularly in the area of functional genomics (Chapter 20). RNAi is ubiquitous in eukaryotes and currently is the preferred tool for investigating the regulation of gene expression in plants, animals, and fungi. This method is becoming increasingly popular because of its wide-ranging applicability in research, which results from compatibility with cell culture and, to a lesser extent, with *in vivo* models. Comparatively speaking, RNAi is much faster and more economical than creating knock-out animals in order to study the function of specific genes. The resulting precision silencing of specific genes also makes RNAi an attractive platform for the discovery and development of life-saving pharmaceuticals. Whereas microarrays are used to correlate an overexpressed or underexpressed gene and a particular disorder, RNAi is used to elucidate a causal link. This technology is so profoundly important that it was voted the 'Breakthrough of the Year' by the prestigious journal *Science*

¹RNA interference is a patented process. Commercial use this process may require licensing from the Carnegie Institution of Washington. For information, contact the Director of Administration and Finance, Carnegie Institution of Washington, 1530 P Street, N.W., Washington, DC 20005. Telephone: 202-939-1118.

in 2002². Not unexpectedly, a number of companies have been founded specifically to provide services and reagents, including premade sets of highly efficient RNA oligonucleotides, to support the exploding market created by this technology.

Essential RNAi nomenclature

Antisense RNA: Single-stranded RNA that is complementary to mRNA and thus has the ability to base pair with it. Like other single-stranded RNAs, antisense RNAs are quite unstable by comparison with dsRNA unless they have been chemically modified.

Argonaute: A highly conserved family of proteins associated with the RNAi process by interaction with small, single-stranded, noncoding RNA, leading to formation of the RNA-induced silencing complex (RISC). It is been suggested that different combinations of argonaute proteins may facilitate RNAi via different pathways and have other roles in the cell (Carmell *et al.*, 2002; reviewed by Hutvagner and Simard, 2008). The Argonaute family proteins are often referred to as the “Ago” proteins.

Dicer: A type III ds-specific RNase that processes long (endogenous or foreign) dsRNA into 21- to 25-bp fragments known as siRNAs through endonucleolytic cleavage. This cytoplasmic enzyme is also able to remove the loop from short hairpin RNA (shRNA), thereby producing short interfering RNA (siRNA). These fragments then become associated with RISC and, if properly designed, hybridize with the complementary sequence of the mRNA to be silenced. Cleavage of precursor microRNA molecules (pre-miRNA) into functional miRNA molecules is also a Dicer enzyme function.

Drosha: A nuclear type III RNase, similar to Dicer, which is involved in the initial maturation step of newly transcribed pri-miRNA (not to be confused with cytoplasmic pre-miRNA).

dsRNA (double-stranded RNA): A duplex consisting of two complementary RNA molecules (sense and antisense) that can be cut by Dicer into siRNAs.

ddRNAi (DNA-directed RNA interference): An *in vivo* method for producing siRNA sequences, under the direction of RNA polymerase III-promoters in mammalian cells. The process involves cell transfection with an expression vector and the resulting siRNAs first pass through the formation of an shRNA intermediate. siRNAs produced *in vivo* are structurally and functionally equivalent to the siRNAs produced *in vitro*.

miRNA (microRNA): Single-stranded, approximately 21 nt RNAs that either (1) block translation of specific mRNAs (translational repression) through mismatched base-pairing; or (2) cause the destruction of specific mRNAs (post-transcriptional regulation) through perfectly matched base-pairing. miRNAs are important for temporal regulation of gene expression during development, cellular differentiation, and myriad housekeeping gene functions. miRNA molecules are transcribed from long 5' capped, polyadenylated transcripts known as pri-miRNAs, and then trimmed by the action of the enzyme Dicer or a closely related family member. Hundreds of different miRNAs have been identified in humans alone.

²Science 298, 2296–2297 (2002).

- PTGS (Post-transcriptional gene silencing):** A term describing the inhibition of gene expression, especially in plants, due to any of a number of causes after a transcript has been produced by a cell. In contemporary literature, the term RNAi has all but replaced PTGS, with which it is synonymous.
- RISC (RNA-induced silencing complex):** A multiprotein complex, including some of the Argonaute family proteins, which associates with the short siRNA fragments produced by the enzyme Dicer. The antisense or “guide” siRNA strand is then used to base pair with the target RNA (mRNA, viral RNA, etc.), thereby marking it for destruction.
- RNAi (RNA interference):** A naturally occurring mechanism for gene silencing induced by the presence of siRNA. Technically, RNAi is PTGS induced by dsRNA!
- shRNA (short hairpin RNA):** Single-stranded RNA molecules possessing sense and antisense domains which facilitate a moderate degree of intramolecular base pairing. The resulting quasi-double-stranded molecule resembles one arm of a typical transfer RNA (tRNA) molecule, with an 8-base single-stranded loop connecting the sense sequence with the antisense sequence. shRNA is a substrate for Dicer, which removes the loop, after which shRNAs effectively become siRNA molecules.
- siRNA (short interfering RNA):** Short, double-stranded RNA molecules, 21 to 23 bp long with 2-nucleotide 3' overhangs (usually a UU dinucleotide). These molecules are at the heart of RNAi functionality.

RNA interference – how it works

RNAi is an endogenous catalytic pathway that is triggered by dsRNA. The trigger can occur naturally, as in the case of a cellular infection by a dsRNA virus, or by the intentional introduction of dsRNA to induce user-directed degradation of the cognate transcript(s). The net result of RNAi is the down-regulation of specific genes by destruction of their mRNA(s).

The notion of inhibiting translation by the creation of double-stranded structures along the length of an mRNA or near its 5' end is not new. As far back as the late 1970s small antisense DNA oligonucleotides were being used to suppress gene expression (Zamecnik and Stephensen, 1978). Antisense RNA was also shown long ago to regulate gene expression in mammalian systems *in vitro* (Nishikura and Murray, 1987), in phages, bacteria, and plants (Green *et al.*, 1986), and in animals (Knecht and Loomis, 1987). Previous attempts to use antisense RNA as a therapeutic agent were thwarted by the fact that antisense RNA is not effective at low concentrations and the fear that using high concentrations of antisense RNA could have cytotoxic effects. After years of using antisense RNA as an *in vitro* inhibitor of gene expression, it was demonstrated (Guo and Kempfues, 1995) that the introduction of *sense* RNA was just as good as antisense RNA for targeted inhibition of gene expression. However, it was subsequently shown (Fire *et al.*, 1998) that using dsRNA was far more efficient than using either antisense RNA or sense RNA. It was later shown that sense RNA caused gene silencing because the sense transcripts were contaminated with antisense transcripts, resulting in the formation of minute

amounts of dsRNA, which is now well known for its potent ability to down-regulate the associated gene.

There are multiple approaches by which RNAi can be induced, each of which has several mechanistic permutations. RNAi is, fundamentally, a two-step process (Fig. 23.1). The first step involves the master enzyme in the RNAi process, a type III endoribonuclease aptly named Dicer. This enzyme, a ubiquitous member of the eukaryotic proteome, is involved in the ATP-dependent digestion of longer dsRNA into 21- to 23bp siRNA molecules with characteristic 3' dinucleotide overhangs on both strands (Zamore *et al.*, 2000). Dicer consists of an amino-terminal helicase domain, two tandem RNase III domains, a PAZ domain (from the names of the proteins PIWI, Argonaute, and Zwiille) which specifically recognizes 3' overhangs, and a carboxy-terminal dsRNA-binding domain (dsRBD). The human-specific family member is known as DICER-1. The long dsRNA can be introduced into the cell (which is sometimes problematic) where it will become a substrate for Dicer; otherwise, recombinant Dicer can be used to generate siRNA *in vitro* after which these short molecules are introduced into the cell by transfection. Dicer also processes pre-miRNA into their mature form.

In the second step, siRNA, regardless of the source, becomes part of a multi-component RNA-induced silencing complex (RISC). RISC consists of endonucleolytic Argonaute proteins and the antisense siRNA strand which guides RISC recognition of the cognate target. Scrutiny of the eukaryotic RISC-associated Argonaute proteins, the number of which is species-specific, reveals four conserved domains: amino terminal domain, the PAZ domain (nucleic acid-binding),

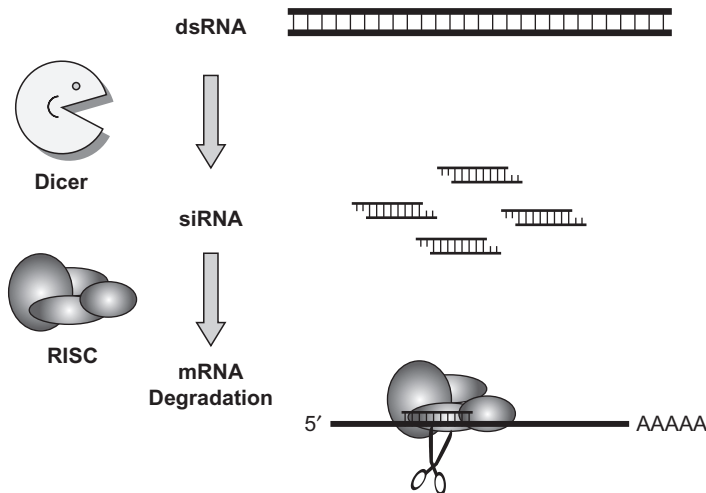


Figure 23.1 Major steps in the RNAi process. dsRNA from any of a number of sources and in any of a number of configurations is cut by the enzyme Dicer into siRNA that, in turn, become part of the multicomponent RISC. This ultimately leads to destruction of the target mRNA and concomitant down-regulation of the associated gene.

the MID-MC domain (5' phosphate- and cap binding³), and the PIWI domain (endoribonuclease activity).

Part of RISC is an ATP-dependent helicase that unwinds the ds-siRNA. The passenger strand, as it is known, is the sense strand of a duplex siRNA that does not associate with the RISC but rather is degraded by it. The antisense component, however, known as the guide strand or 'guide RNA', is able to base-pair with the mRNA to be silenced. The formation of a double-stranded region between the antisense component of the silencing RNA and the cognate mRNA dooms it. RISC cuts the mRNA between nts 10 and 11 and does so in a non-ATP-dependent manner (Elbashir *et al.*, 2001b). The mRNA is then further degraded, preventing any level of interaction with the cellular translation machinery. The antisense guide RNA remains associated with the RISC, thereby sustaining its ability to direct the cleavage of other identical target mRNAs. As expected, transfection of cells with only the sense siRNA strand elicits no silencing, and transfection with the antisense siRNA sometimes produces a very low level of silencing. In contrast, transfection with ds-siRNA produces a silencing response which is much greater, to the extent that synthesis of the encoded protein can be suppressed altogether.

The three fundamental methods which are used to elicit gene silencing are (1) siRNA; (2) shRNA; and (3) miRNA. There are advantages and disadvantages to the siRNA and shRNA approaches (Table 23.1), though it is difficult to definitively select one method as clearly superior because RNAi is so new and is changing rapidly. It can be helpful to attempt the same experiment more than one way, and then decide which approach is best suited for the business at hand. Each of these approaches has permutations, though they all have the ability to target specific mRNAs for scission.

Table 23.1 Comparison of siRNA and shRNA for RNAi

siRNA	shRNA
<ul style="list-style-type: none"> • Expensive • Easier to design than shRNA • Produced by chemical synthesis or by <i>in vitro</i> transcription and processing • Toxicity often an issue, especially in primary cells and neurons • Suppression effect lasts for 5–7 days • Less time from start to finish • Focuses on short-term effects 	<ul style="list-style-type: none"> • Relatively economical • More difficult to design than siRNA • Primarily expression-based <i>in vivo</i> or <i>in vitro</i>, but can be produced by chemical synthesis • Toxicity usually not a problem • Can be used to generate stable cell lines • Sustained suppression of gene expression • More time from start to finish • Focuses on long-term effects

³The 5' cap-binding ability of some of the Argonaute proteins suggests an important regulatory role at the translational level because one function of the cap is identification of the molecule as an mRNA and subsequent assembly of the translation apparatus.

Before discussing the methods of RNAi induction, it is worth noting that there is much confusion as to the functional similarities and differences between siRNA and miRNA. In terms of the net result, miRNA and siRNA actions are often indistinguishable. Thus, the major difference is not what these molecules do but rather where they come from. miRNAs result from endogenous transcription of genomic DNA within the cell while siRNAs are the result of either the processing of long dsRNA, the processing of shRNA that are produced by expression vector transcription, or by direct introduction via transfection. Both siRNAs and miRNAs mature through the action of an enzyme known as Dicer, though different Dicer-type enzymes support the miRNA- and siRNA pathways in many organisms. In either case, both types of small RNAs become associated with a RISC-like complex. The major functions of miRNA are translational repression and mRNA cleavage, while siRNAs tend to be associated with mRNA cleavage alone.

siRNA approach

siRNA molecules can be easily produced *in vitro* and then introduced into target cells by transfection. For example, individual sense and antisense RNA oligos can be synthesized and then annealed. An alternative strategy is to synthesize a 500- to 750bp dsRNA and then digest it with Dicer to create a pool of siRNA molecules (Fig. 23.2). The entire siRNA pool can then be transfected into the cell to induce RNAi. This approach may be particularly beneficial when the precise siRNA sequence needed to silence a gene is unclear. By providing extensive coverage of a target mRNA with a complex siRNA pool, it is likely that at least some of the siRNA members will be able to direct the destruction of that target. This is essentially the brute-force approach to attacking the problem.

Long dsRNA can also be introduced into the cell through transfection using cationic liposomes or pinocytotic uptake (Gruber *et al.*, 2004). The length of dsRNA is of critical importance in experiments involving mammalian cells because dsRNA longer than 30 bp will induce an interferon response (Martinand *et al.*, 1998) that is likely to result in apoptosis. In the cell, siRNA is produced when perfectly paired dsRNA is processed by the enzyme Dicer, producing the characteristic 21- to 23 bp siRNAs, a length has been observed ubiquitously. It has been suggested that any dsRNA that is 26–30 nt long will most likely induce RNAi (Parrish *et al.*, 2000). Many investigators prefer to transfect siRNA produced *in vitro* into mammalian cells, however, rather than introducing long dsRNA into the cell (Elbashir *et al.*, 2001a; Elbashir *et al.*, 2001b) in order to bypass the potential induction of the interferon response. In general, the siRNA approach described here is best-suited for studying the short-term effects of gene silencing.

Another strategy involves siRNA production directly in the cell via transcription of the (+) and (–) strands of a previously transfected gene in a suitably configured expression vector (Brummelkamp *et al.*, 2002). This development was a

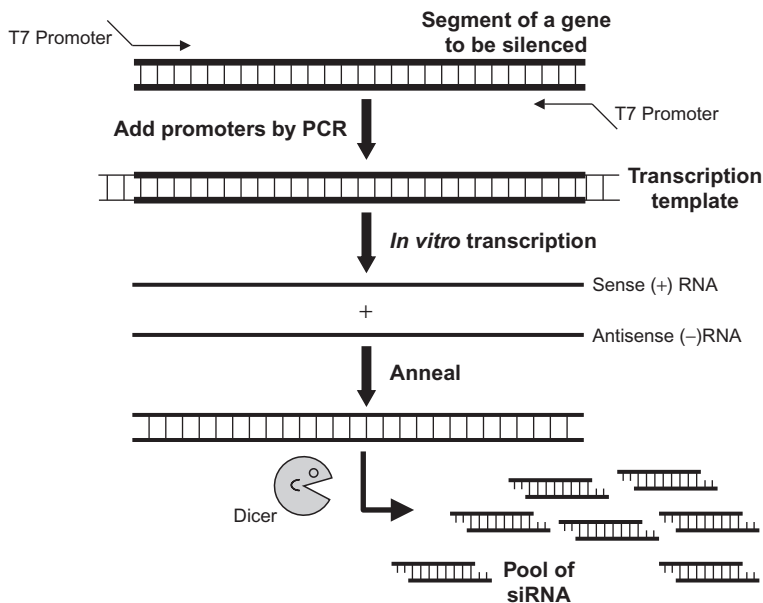


Figure 23.2 *In vitro* synthesis of a heterogeneous pool of siRNA. T7 transcription promoters are added by PCR (described in Chapter 18). The PCR product is then used to support *in vitro* transcription, thereby producing both sense and antisense RNA, because identical promoters are found at each end of the transcription template. After allowing several hours for annealing to occur, a pool of siRNA is generated by Dicer. The siRNA molecules are then able to interact with RISC, as usual, for RNAi induction.

defining moment in the short history of RNAi because it represented a method for long-term silencing of specific genes and is described in the following section.

shRNA approach

shRNAs (Paddison *et al.*, 2002; Yu *et al.*, 2002) are an alternative to the use of dsRNA molecules or siRNAs used to suppress gene expression. shRNAs are single-stranded RNA molecules that, by virtue of inverted repeats exhibit intramolecular base-pairing. Once formed, the shRNA molecules are processed by Dicer and are subsequently able to cause gene suppression, much as occurs through the use of siRNA.

shRNAs may be produced *in vitro* (Fig. 23.3) but are most often produced *in vivo* through the use of transcription vectors and under the direction of RNA polymerase III. shRNA molecules are characterized by a sense (+) component consisting of 21 to 29 bases that are exactly the same sequence as the mRNA to be silenced, an 8-base single-stranded loop, and then an antisense (-) sequence that is precisely complementary to the sense sequence. The siRNA is also characterized by a terminal UU dinucleotide at the 3' end (Fig. 23.4). If produced *in vitro*, shRNAs may be introduced into mammalian cells via liposome

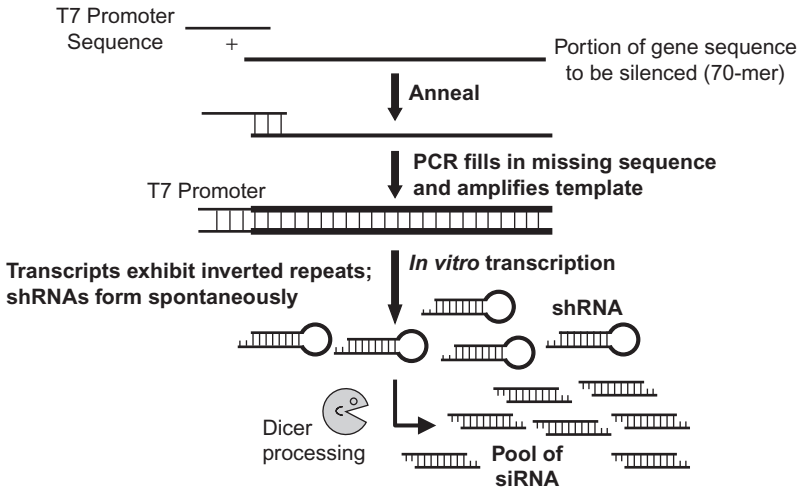


Figure 23.3 *In vitro* method for production of shRNA. PCR is used to construct a transcription template. In contrast to the method shown in Figure 23.2, a T7 promoter is added only to one side of the template. The template is constructed in such a way that it contains inverted repeats consisting of both sense (+) and antisense (–) domains. The structure of the resulting transcripts favors rapid formation of shRNA molecules. Subsequent cleavage by Dicer prepares a pool of siRNA that can induce RNAi.

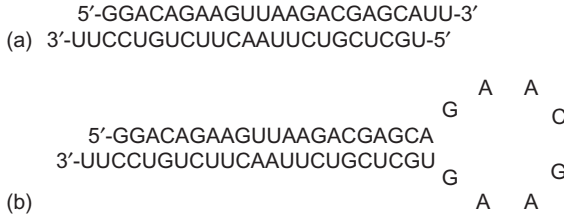


Figure 23.4 Comparative anatomy of siRNA and shRNA. (a). siRNA molecules consist of two complementary RNA molecules, characterized by a 3'-terminal UU dinucleotide overhang on each strand. (b). shRNA molecules consist of one, single-stranded molecule. Intramolecular sense and antisense domains favor the rapid formation of a hairpin-like structure. The folded molecule has an 8-base, single-stranded loop and, like siRNA molecules, exhibits a 3' UU overhang. The loop joining the sense and antisense domains is subsequently removed by Dicer, at which point the shRNA essentially becomes a siRNA. Note that both of these molecules would silence the same mRNA, because they exhibit exactly the same base composition in their respective double-stranded regions.

transfer though, like many other RNAi applications, delivery continues to remain a problem.

In the case of *in vivo* synthesis, the transcription vector may be introduced into the cell by any standard transfection technique, though lentivirus vectors are known for high-efficiency transfection of both dividing and quiescent cells, highly differentiated cells such as neurons, and primary cells. This approach is

often referred to as DNA-directed RNA interference, or simply ddRNAi. This method occasionally induces the interferon response in mammalian cells, similar to that which occurs when long dsRNA enters from the outside and which the cell perceives as a viral assault. Some scientists use this approach, however, because it represents a significant advantages such as cost savings, long-term interference, and the ability to turn the RNAi mechanism on and off at will via inducible promoters.

Transcription vectors used for this purpose contain inverted repeat sequences such that newly transcribed RNA will form a hairpin structure that, as a substrate for Dicer, will be processed into siRNA (Fig. 23.5). It is now commonplace to combine several shRNA cassettes into a single vector. This technology emerged as a consequence of the need to knock down the expression of more than one gene associated with a particular pathway. At some point, it may be possible to produce transgenic animals which contain these types of constructions, but only insofar as the encoded shRNA does not target genes that are required for fertility or viability.

siRNA delivery methods into mammalian cells

Needless to say, both medical and economic benefits will result from the production of an RNAi drug. While the science behind RNA silencing is fairly well

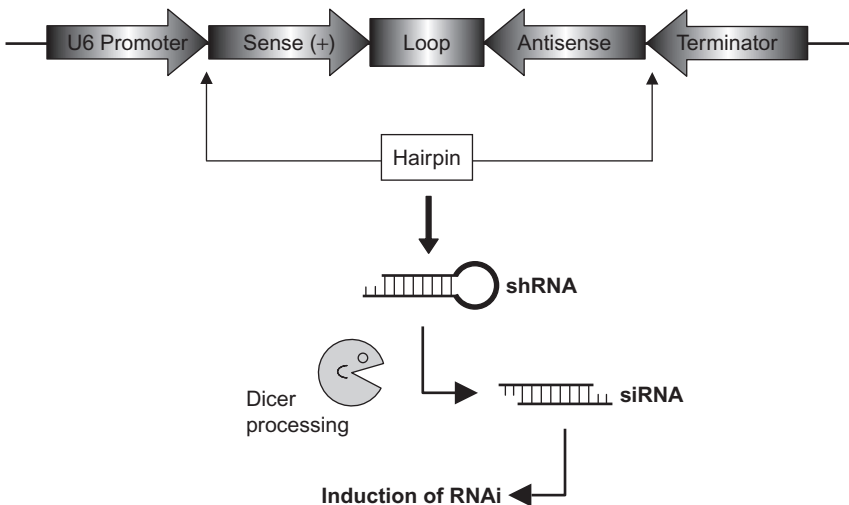


Figure 23.5 *In vivo* transcription template for the production of shRNA. The construction is under the influence of a U6 promoter (RNA polymerase III). Notice that the sense and antisense domains are oriented facing each other, separated by a short region encoding what will become the single-stranded loop. The resultant shRNA transcripts are substrates for Dicer, which will “promote” them to siRNA and induce the RNA interference pathway. In plants, a useful promoter for driving RNAi via this approach is the efficient constitutive cauliflower mosaic virus 35S promoter.

understood, a major impediment is delivery of the molecules into the cell. Much effort has been invested in attempts to optimize liposome delivery of RNAi molecules and while this approach works fairly well *in vitro*, there are toxicity issues *in vivo* associated both with the delivery as well as a consequence of the modified chemistry of the oligonucleotides. In general, the negative charge density associated with all polynucleotides makes it difficult for these molecules to move through the hydrophobic interior of the cell's lipid bilayer.

It is clear that a number of approaches can result in the suppression of gene expression by RNAi. Often, the choice of one approach over another is based on the delivery method. Low transfection efficiency is the primary cause of poor- or altogether unsuccessful RNA silencing. For example, if a particular cell population does not transfect efficiently, then shRNA produced *in vivo* may be the answer. For each new cell type, it is necessary to optimize the transfection procedure and the reagents selected in order to promote successful silencing.

dsRNA of any length can be produced *in vitro* and delivered into the cell by transfection. Likewise, the delivery of siRNA and shRNA can occur in this manner. As mentioned above, the problem in mammalian cells (but not other cell types) is that any dsRNA greater than 30bp long will stress the cell, eliciting an interferon-mediated response. To preclude this difficulty, shRNA can be produced *in vivo*, because dsRNA produced in the intracellular milieu does not evoke the interferon response. *In vivo* transcription strategies involve the design of an inverted repeat that favors rapid intramolecular transcript folding and subsequent initiation of the RNAi pathway. The efficiency of gene silencing must ultimately be validated by Western analysis, other immunological techniques, and/or by RT-PCR.

miRNA

The understanding of gene regulation, at least in the traditional sense, was turned upside-down by the discovery of miRNAs (Lee *et al.*, 1993). miRNAs are endogenous 19- to 25 nt regulatory molecules that are transcribed by RNA polymerase II⁴. The primary product of transcription is a long, single-stranded RNA which folds over on itself to produce a single-stranded loop, a stem consisting of about 20 base-pairs, and unpaired sequences dangling from the 5' and 3' ends. This precursor, known as pri-miRNA is trimmed by the nuclear enzyme Drosha, the action of which removes the non-base-paired 5' and 3' ends. This produces a pre-miRNA which is then exported into the cytoplasm where the enzyme Dicer acts to remove the single-stranded loop from what is known as pre-miRNA. The result of this later cleavage event is short double-stranded miRNA (Lee *et al.*, 2002), which is sometimes designated miRNA-miRNA*, with the asterisk representing the passenger strand that does not become associated with the RISC and which, in fact, will be degraded by it. Originally believed only to suppress protein synthesis by base-pairing to the

⁴miRNAs can also be produced from introns.

cognate RNA, miRNAs are now known to have the ability to promote protein synthesis (Vasudevan *et al.*, 2007).

miRNAs are known to govern cellular- and organism physiology at the levels of cell proliferation, differentiation, cellular senescence, apoptosis, and the progression of the disease state. miRNAs almost always base-pair (imperfectly) with the target mRNA in animal cells and manage to suppress translation; in plant cells, however, precision base-pairing directs the miRNA-mediated destruction of the mRNA. Thousands of miRNAs have already been identified in many organisms and have been added to the rapidly expanding miRNA database: <http://microrna.sanger.ac.uk/>.

Effective design of siRNAs

Without a doubt, the field of functional genomics has become supercharged by RNAi technology. As such RNAi has become the number one method for examining gene function *in vivo* and *in vitro*, and the success of RNAi is dependent on the proper design of dsRNA. Now that the human genome has been sequenced, bioinformatics is indispensable in the process of RNA oligonucleotide design. In human cells, a single-mismatch will greatly compromise RNAi, and more than one mismatch will inhibit RNAi altogether. Conveniently, companies that sell RNAi-related products generally provide free web-accessible design tools (Table 23.2).

At its most fundamental level, dsRNA must inactivate the intended mRNA. It is the sequence of the dsRNA that dictates precisely where the mRNA will be cleaved. It is very important that the dsRNA does not perturb the expression of any other genes because only great specificity and the absence of “side-effects” will be able to unlock the function of specific genes and the physiology of the cell. It is widely believed that all genes can be silenced by RNAi, although it is abundantly clear that not all siRNAs are capable of doing that job.

There are basic rules to which one should adhere in rational siRNA design, and most companies involved in RNAi technology have proprietary algorithms

Table 23.2 Selected siRNA Design Tools

Exon-Intron Database	http://hsc.utoledo.edu/bioinfo/eid/
The RNAi Web	www.rnaiweb.com/RNAi/siRNA_Design/
Promega	www.promega.com/siRNADesigner
Qiagen	www.qiagen.com/siRNA
Dharmacon	www.dharmacon.com/DesignCenter
Ambion	http://ambion.com/techlib/misc/siRNA_tools.html
Oligo Engine	www.oligoengine.com
Whitehead Institute	http://jura.wi.mit.edu/bioc/siRNAext

that will maximize the probability of effective gene silencing. Some of the basic siRNA design principles are intuitive and reminiscent of the standard guidelines for designing oligonucleotides for PCR applications. A comprehensive overview of siRNA design basics (informally called *Tuschl's Rules*) can be accessed online (www.rockefeller.edu/labheads/tuschl/sirna.html). Briefly, these and other rules suggest that one should limit the GC content to 50% or less (or at least keep it within the 40–60% range) and endeavor to identify targets that exhibit at least a three- or four-base difference from any other similar-sized sequence in the genome (usually a 21-base sequence). As with PCR primers, one should avoid areas that are likely to form undesired secondary structures so as to maintain access accessibility of target sequence. Canonical base-pairing between nts 2–10 of the miRNA and the target is also important, and siRNAs should be designed with a favorable (low) free energy associated with siRNA-mRNA duplex formation. Effective siRNAs are 21- to 23 nt duplexes exhibiting symmetrical UU 3' overhangs and the target site and the mRNA should begin with AA, and targeting a conserved region of the mRNA away from the 5' end of the molecule seems to work best. When planning to perform ddRNAi *in vitro* or *in vivo*, it is very important to avoid the inclusion of four or more of the same base in a row, which will cause premature termination of transcription initiated from the commonly utilized human U6 promoter (Paul *et al.*, 2002). Using a stretch of four to five thymidines is how transcription termination is encoded in the ddRNAi-related methods. Finally, obviously, and perhaps most importantly, one should perform a BLAST search to determine the extent to which an siRNA will interact with mRNAs other than that which is intended.

Because rates of suppression vary among transcripts, it is neither unusual nor unreasonable to test three or more siRNAs to find at least one sequence that will produce a profound knockdown in target gene expression (>90%) by targeting different areas of the mRNA to ensure its silencing. It is usually a good idea to demonstrate silencing through the use of two separate siRNAs to demonstrate specificity. Post-validation, one may wish to try using both siRNAs simultaneously in an attempt to maximize the silencing effect. Alternatively, synthesis of very large numbers of siRNAs, representing the entire coding region of a transcript, is often an effective means for ensuring silencing. This can be performed beginning with a very long dsRNA which will then be processed by Dicer to produce a heterogeneous pool of siRNAs capable of binding along most of the length of the transcript.

miRNA efficiency is also species-dependent. For example, most plant miRNAs base-pair to their mRNA targets with very few, if any, mismatches. In animals, however, imperfect base-pairing is far more common and the parameters governing translational control are somewhat unclear. In the event that an siRNA does not function as expected, especially using a cell culture model, it is worth taking the time to verify identity of the cells. As bizarre as it may seem, there have been many situations in which cell cultures have unwittingly become contaminated with and then overrun by other cells (e.g., HeLa cells), often from a different species.

Finally, the traditional chemical method for RNA oligonucleotide synthesis suffers from an overall lack of coupling efficiency with respect to the length of the molecules produced and achievable level of purity. The RNAi revolution has benefitted from a novel method known as 5'-silyl-2'-acetoxy ethyl orthoester chemistry (Scaringe *et al.*, 1998), or 2'-ACE™. Succinctly, this method is based on a novel protecting group scheme and which is superior in terms of speed, efficiency, and product purity and stability during shipping. Upon arrival the RNAs can be easily deprotected in aqueous solutions and, if desired, the individual RNAs can be deprotected and annealed to form siRNA simultaneously. This latter approach imparts an added level of stability. On a related note, and in an attempt to improve the efficiency of RNAi, the use of 2'-F-CTP and 2'-F-UTP (fluorine in the pentose 2' position normally occupied by oxygen) in the synthesis of dsRNA greatly stabilizes the RNA, making it resistant to many of the commonly observed nucleases found in human cells and on human hands (Capodici *et al.*, 2002; Kariko *et al.*, 2004).

RNAi and alternative transcript splicing

The impressive size of the transcriptome is a result of the alternative splicing of transcripts from a single locus as well as the potential for alternative transcription start sites (TSS) in a temporal- or in a tissue-specific manner. Even the most rationally designed siRNAs can be duds, perhaps due to the plethora of alternative splicing events that characterize the mRNAs produced from many, if not all, higher eukaryotic loci. For example, if a siRNA is directed toward an exon that is spliced out during mRNA maturation (Fig. 23.6), then that particular siRNA will not be able to cause gene silencing. As it is not always possible to predict the transcriptional behavior of a genetic locus or the posttranscriptional behavior of the resulting transcript, designing several different siRNAs per target is a good idea, not to mention the overall unpredictability of any newly designed siRNA.

For the very same reasons, one should avoid targeting introns unless, for example, it is known that a particular tissue maintains a “cryptic intron” as a consequence of more than one functional TATAA-box element in the associated promoter (Bassett *et al.*, 2009). Similarly, one should avoid targeting known introns and SNPs (single nucleotide polymorphisms) because a single mismatch between mRNA and miRNA is all that is necessary to change the outcome of an experiment drastically.

Often, siRNAs are designed that target regions of the mRNA at least 50–75 nt downstream from the initiation (AUG) codon and at least 50–75 nt upstream from the 5' end of the poly(A) tail or at least that distance upstream from the 3' end of non-poly(A) transcripts. That RNA binding proteins already associated with target mRNA may reduce the binding potential for RISC is the reason why most investigators tend to stay away from the ends of these molecules.

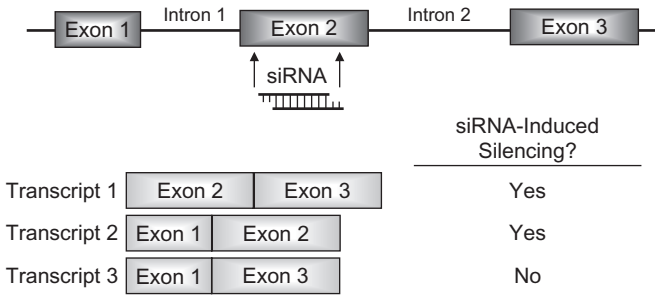


Figure 23.6 Differential splicing can affect the success of gene silencing by RNAi. Alternative splicing of hnRNA is a normal part of mRNA biogenesis for many genes. If siRNA is target toward an exon that has been spliced out of the mRNA precursor, RNAi will not be induced because of an inability of the siRNA to associate with the template. A similar consequence may result from the use of an alternative transcription start site. Splicing patterns and the precise location of transcription initiation may well be principal reasons why many well-designed siRNAs fail to suppress the expression of some genes.

In vitro and *in vivo* issues

Among the biggest problems associated with RNAi functionality is the resistance of cells grown in suspension culture and primary cells to siRNA uptake, compared to anchorage-dependent cell lines. Keep in mind that the degree of confluency in a tissue culture vessel as well as the passage number (associated with the number of cell population doublings), may both impact transfection efficiency. For example, cell populations with low passage numbers and which are about 50% confluent when exposed to the transfection complex usually produce superior results.

Some of the toxicity difficulties associated with classical CaPO₄ transfection were assuaged years ago by the development of a plethora of liposome delivery vehicles, although *in vivo* inflammatory responses, and *in vitro* cytotoxicity, are sometimes noted in response to the use of lipid transfection products. Even when these liposome preparations don't kill cells outright, it has been noted that liposome transfection may alter the transcription file of the cell or even activate the apoptotic pathway (Sledz *et al.*, 2003). Overcoming these obstacles is of paramount importance if the application of the technology is to progress to the next clinical level.

Methods will also need to be developed that allow efficient introduction of RNAi tools into primary cells, since these most closely mirror the normal biochemistry of the cell. Let's face it...cell lines have something wrong with them and tend to experience significant changes in the architecture and composition of their genome as they are continually passaged in culture. There are some mammalian cells, neurons and primary cells in particular, that are naturally resilient to transfection, often mandating nearly heroic efforts to achieve even low transfection efficiency. Interestingly, in some cases, siRNA seems to be

taken up into some cells more efficiently than double-stranded plasmid DNA (Krichevsky and Kosik, 2002), opening the door for more a more structured, in-depth analysis of gene expression in cell types that have been resistant to this type of analysis to date. Novel methods for improving transfection efficiency (Weil *et al.*, 2002) are now being explored for the introduction of plasmids, dsRNA, siRNA, and a plethora of other potentially useful compounds, including the use of cholesterol-conjugate siRNAs.

Sometimes, gene redundancy may hamper discovering the true function of a gene. Gene duplication or the need the knock down expression of more than one gene associated with a biochemical pathway may put a damper on productivity. In order to address this potentially serious difficulty, one should consider expression vector constructions encoding more than one shRNA. Recently described vectors (Steuber-Buchberger *et al.*, 2008) have been engineered for this purpose.

The overall RNAi efficiency in an experiment is also a function of the abundance and the stability of both the target mRNA and the encoded protein, and the precise region of the mRNA targeted for scission. Recall also that siRNA and mRNA need to base-pair with 100% complementarity because mismatches between these molecules are known to reduce the RNAi effect dramatically in mammalian cells, presumably by inducing an interferon response. This has been clearly demonstrated over the years by the use of poly I:C, a synthetic double-stranded RNA polymer used for induction and characterization of the interferon stress response⁵.

Another issue associated with the use of siRNAs for *in vivo* applications is the presence of cytotoxic compounds that are associated with their *in vitro* synthesis. In particular, the removal of trace heavy metals and the maintenance of an endotoxin-free environment are of paramount importance. Often, investigators who are unfamiliar with the aseptic measures needed to maintain cells in culture or when working with animals, are surprised by the magnitude of contamination problems that can result and the ensuing inflammatory response in animals, respectively. In the case of cells maintained in culture, lipopolysaccharide (LPS; classic endotoxin) may trigger the up- or downregulation of genes that might not otherwise be associated with the investigation at hand. Surprising, even aliquots of allegedly sterile water can be a source of LPS if the sterility of the aliquot has been compromised. Several laboratories and reagent suppliers test each product lot, often using the standard Limulus amoebocyte lysate (LAL) assay, to ensure that endotoxin levels are exquisitely low levels or absent altogether. Succinctly, one should seek guidance from someone who is familiar with standard aseptic techniques associated with the propagation of eukaryotic cells before moving forward.

⁵Ampligen® (Hemispherx Biopharma; Philadelphia, PA), is an experimental drug based on the ability of dsRNA to induce an interferon response. Unlike poly I:C, however, Ampligen is a homopolynucleotide duplex containing a mismatched base analogue (poly (I)-Poly (C-U)) (Strayer *et al.*, 1982). Visit www.Hemispherx.net for details.

Table 23.3 RNAi Advantages and Disadvantages

Advantages	Disadvantages
Examine function(s) of one gene at a time	Cost: RNA oligos are more expensive than DNA oligos
Highly specific to the gene(s) of interest	Extensive training of technicians is required
Systematic approach to investigate gene function and which genes are involved in which pathways	dsRNA >30 bp upregulates interferon response genes in mammalian cells, leading to apoptosis
Highly evolutionarily conserved process more efficient at mRNA destruction than ribozymes	Some cells, especially neurons and primary cells, tend to be resistant to RNAi
Much less toxic than phosphorothioate-based oligonucleotides	Reproducibility is sometimes a problem
Easy to deliver to mammalian cells	Some siRNAs, no matter how rationally designed, simply do not work, mandating the trial-and-error approach
Well suited for drug target identification	Interlaboratory comparisons often difficult because of variations in protocols
Potent: dsRNA is effective at much lower concentrations than antisense oligos	Synthetic RNAi oligos need to be extremely pure (desalting, deprotection, HPLC purification, mass spec, etc.)

It is ironic that even though RNAi appears to be widespread throughout the plant, animal, and fungi kingdoms, this powerful regulatory mechanism utilizing dsRNA is poorly understood. While there appear to be several common denominators in the RNAi process from one eukaryotic organism to the next, it is becoming increasingly clear that there are subtle differences. RNAi is in its early years for sure, and each new development in the technology requires careful evaluation in the model system for which its use is intended. Some of the advantages and disadvantages associated with the use of this still-novel technology are presented in Table 23.3. RNAi as a tool is currently a long way from perfection, but it is definitely attractive for the study of gene function. If gene silencing fails completely, it is probably time to design new siRNAs. Another appropriate action is to check sequencing data and to *consider its source*.

RNAi validation

In order to claim a “silencing victory” in a given experiment, it is necessary to demonstrate that the flow of genetic information was disrupted. There are three common methods for demonstrating that posttranscriptional silencing has occurred. These methods involve the direct assay of the abundance of the mRNA, the protein, or both.

RT-PCR approaches

Without a doubt, direct mRNA assay by RT-PCR is a sure way to determine if the abundance of a transcript has increased, decreased, or remained unchanged. As described extensively in Chapter 18, the RNA is purified from the cells under investigation, reverse-transcribed in cDNA, and then amplified by PCR with a suitable set of primers. If a particular transcript is being targeted for RISC-mediated cleavage, its abundance in the cell is expected to decrease, sometimes dramatically, and in a manner directly associated with the efficiency of silencing. A logical way by which to approach RT-PCR for this specific application is to design at least two sets of primers for each transcript under investigation. The first and most obvious set of primers would span the region where mRNA cleavage would be expected; cDNA synthesis cannot extend beyond the cleavage point and, therefore, the 5' primer would be unable to participate in amplification by PCR because of the truncated nature of the 3' end of the first-strand cDNA. The second set of primers should anneal to the cDNA in a region corresponding to the 3' end of the cDNA in order to demonstrate that the cDNA had, in fact been synthesized and its integrity compromised by the mechanics of RNAi.

Northern analysis

RT-PCR is a fine quantitative tool for the rapid assay of transcriptional activity associated with a particular gene locus. However, the PCR product represents only a small section of the mRNA, rather than the full-length transcript. Some investigators prefer to demonstrate successful RNAi through the assay of total cellular or total cytoplasmic RNA by Northern analysis so as to be confident that the size of the transcript species targeted for destruction is consistent with known size of that transcript. Of course, successful Northern analysis is predicated upon having a sufficiently abundant transcript to assay, at least in the control samples, because Northern analysis is a rather low-sensitivity assay (Chapter 11). If using mRNA abundance to monitor gene knockdown efficiency, be sure to measure the abundance of a non-target mRNA, too, as a suitable control.

Western analysis

The most direct means by which to demonstrate a change in the abundance of a protein is to assay it directly. Specific antibodies are an effective means by which to monitor the reduction of a protein when the cognate mRNA is targeted for destruction via RNAi. It is also possible to monitor changes in the quantity and subcellular location of a protein by *in situ* immunofluorescence. This method is also used to monitor overall transfection efficiency.

While the mechanics of RNAi are undoubtedly initiated immediately after the appearance of siRNA in the cell, the ability to measure gene knockdown, particularly at the level of protein synthesis may not be apparent for a day or two; if the protein of interest has a long half-life, even longer periods may be

necessary. If protein knockdown is not observed, one may wish to assay the abundance of the corresponding mRNA either by RT-PCR or, in the case of abundant transcripts, by Northern analysis.

Applications

The immense potential for RNAi has yet to be tapped, primarily because the scientific community is just now learning about novel regulatory pathways in the cell. There have been several exciting recent developments in plants and in animals involving siRNA and miRNA, suggesting methods by which these molecules might be controlled in order to provide some type of health or economic benefit. For example:

1. It has recently been discovered that the control of cellulose production in barley is controlled by an siRNA (Held *et al.*, 2008). By interfering with this siRNA-mediated downregulation of the cellulose synthesis pathway, it should be possible to induce enhanced cellulose production in certain plant species for biofuel production.
2. Also in plants, *trans*-acting small interfering RNAs (tasiRNA; Dunoyer *et al.*, 2007) have been discovered. Briefly, miRNA mediates the cleavage of a single-stranded pre-siRNA transcript in the vicinity of the 3' end. The resulting cleavage product is converted into dsRNA by the action of an RNA-dependent RNA polymerase which, in turn, becomes a Dicer substrate and from which siRNA(s) are produced. Finally, the siRNA becomes RISC-associated, thereby directing the scission of the target mRNA. Another type of siRNA, rasiRNA (repeat-associated small interfering RNA) is highly abundant in germline cells and has been implicated in histone modification. Understanding these gene regulatory circuits may provide additional hints as to how one might gain control of these pathways.
3. Over-expression of miRNA 182 in human metastatic melanoma has recently been documented (Segura *et al.*, 2009), and appears to pushing cells through the cell cycle much like an activated oncogene. Regaining control of this molecule, either at the level of its synthesis or its destruction, could prove to be a valuable weapon against this deadly disease.

The use of RNAi has profound ramifications in many facets of cell biology. Through the ability to target for study of one gene at a time, RNAi has become another mainstream tool of the molecular biologist. Its ubiquity, functionality, and compatibility with basic- and applied research make RNAi an attractive platform for addressing many problems in the life sciences.

References

- Bassett, C.L., M.E. Wisniewski, T.S. Artlip, G.A. Richart, J.L. Norelli and R.E. Farrell, Jr. (2009). Comparative expression and transcript initiation of three peach dehydrin genes. *Planta* 20, 107–118.
- Brummelkamp, T.R., R. Bernards and R. Agami. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550–553.

- Capodici, J., K. Kariko and D. Weissman. (2002). Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference. *J. Immunol.* **169**, 5196–5201.
- Carmell, M.A., Z. Xuan, M.Q. Zhang and G.J. Hannon. (2002). The argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* **16**, 2733–2742.
- Dunoyer, P., C. Himber, V. Ruiz-Ferrer, A. Alioua and O. Voinnet. (2007). Intra- and intercellular RNA interference in *Arabidopsis thaliana* requires components of the microRNA and heterochromatic silencing pathways. *Nat. Genet.* **39**, 848–856.
- Elbashir, S.M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498.
- Elbashir, S.M., W. Lendeckel and T. Tuschl. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200.
- Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver and C.C. Mello. (1998). Potent and specific interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
- Green, P., O. Pines and M. Inouye. (1986). The role of antisense RNA in gene regulation. *Ann. Rev. Biochem.* **55**, 569–597.
- Gruber, J., G. Boese, T. Tuschl, M. Osburn and K. Weber. (2004). RNA interference by osmotic lysis of pinosomes: liposome-independent transfection of siRNAs into mammalian cells. *BioTechniques* **37**(1), 96–102.
- Guo, S. and K.J. Kemphues. (1995). *Par-1*, a gene required for establishing polarity in *C. elegans*, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81**(4), 611–620.
- Hammond, S.M., E. Bernstein, D. Beach and G.J. Hannon. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293–296.
- Held, M.A., B. Penning, A.S. Brandt, S.A. Kessans, W. Yong, S.R. Scofield and N.C. Carpita. (2008). Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley. *Proc. Natl. Acad. Sci.* **105**, 20534–20539.
- Hutvagner, G. and M.J. Simard. (2008). Argonaute proteins: key players in RNA silencing. *Nat. Rev.* **9**, 22–32.
- Jankowski, B.A., S.T. Younger, D.B. Hardy, R. Ram, K.E. Huffman and D.R. Corey. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* **3**, 166–173.
- Kariko, K., P. Bhuyan, J. Capodici and D. Weissman. (2004). Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *J. Immunol.* **172**, 6545–6549.
- Knecht, D.A. and W.F. Loomis. (1987). Antisense RNA inactivation of myosin heavy chain gene expression in *Dictyostelium discoideum*. *Science* **236**, 1081–1086.
- Krichevsky, A.M. and K.S. Kosik. (2002). RNAi functions in cultured mammalian neurons. *Proc. Natl. Acad. Sci.* **99**, 11926–11929.
- Lee, R.C., R.L. Feinbaum and V. Ambros. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.
- Lee, Y., K. Jeon, J.T. Lee, S. Kim and V.N. Kim. (2002). microRNA maturation: step-wise processing and subcellular localization. *EMBO J.* **21**, 4663–4670.

- Li, L.C., S.T. Okino, H. Zhao, D. Pokot, R.F. Place and R. Dahiya. (2006). Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci.* **103**, 17337–17342.
- Llave, C., K.D. Kasschau and J.C. Carrington. (2000). Virus-encoded suppressor of post-transcriptional gene silencing targets a maintenance step in the silencing pathway. *Proc. Natl. Acad. Sci.* **97**, 13401–13406.
- Martinand, C., T. Salehzada, M. Silhol, B. Lebleu and C. Bisbol. (1998). The RNase L inhibitor (RLI) is induced by double-stranded RNA. *J. Interf. Cytok. Res.* **18**, 1031–1038.
- Merrett, A.F., W. Aufsatz, J. van Der Winden, M.A. Matzke and A.J. Matzke. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**, 5194–5201.
- Mullis, K.B. and F.A. Faloona. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335–350.
- Mullis, K.B., F.A. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Sym. Quant. Biol.* **51**, 263–273.
- Napoli, C., C. Lemieux and R. Jorgensen. (1990). Introduction of a chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. *Plant Cell* **2**, 279–289.
- Nishikura, K. and J.M. Murray. (1987). Antisense RNA of proto-oncogene *c-fos* blocks renewed growth of quiescent 3T3 cells. *Mol. Cell. Biol.* **7**, 639–649.
- Paddison, P.J., A.A. Caudy, E. Bernstein, G.J. Hannon and D.S. Conklin. (2002). Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16**, 948–958.
- Parrish, S., J. Fleenor, S. Xu, C. Mello and A. Fire. (2000). Functional anatomy of a dsRNA trigger: differential requirement of the two trigger strands in RNA interference. *Mol. Cell* **6**, 1077–1087.
- Paul, C.P., P.D. Good, I. Winer and D.R. Engelke. (2002). Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**, 505–508.
- T.Pélissier and M. Wessnegger. (2000). A DNA target of 30 bp is sufficient for RNA-directed DNA methylation. *RNA* **6**, 55–65.
- Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich and N. Arnheim. (1985). Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354.
- Scaringe, S.A., F.E. Wincott and M.H. Caruthers. (1998). Novel RNA synthesis method using 5'-silyl-2'-orthoester protecting groups. *J. Am. Chem. Soc.* **120**, 11820–11821.
- Segura, M.F., D. Hanniford, S. Menendez, L. Reavie, X. Zou, S. Alvarez-Diaz, *et al.* (2009). Aberrant miR-182 expression promotes melanoma metastasis by repressing FOXO3 and microphthalmia-associated transcription factor. *Proc. Natl. Acad. Sci.* **106**, 1814–1819.
- Sledz, C.A., M. Holko, M.J. de Veer, R.H. Silverman and B.R. Williams. (2003). Activation of the interferon system by short-interfering RNAs. *Nat. Cell Biol.* **5**, 834–839.
- Strayer, D.C., W.A. Carter, I. Brodsky, D.H. Gillespie, J.J. Greene and P.O. Ts'o. (1982). Clinical studies with mismatched double-stranded RNA. *Tex. Rep. Biol. Med.* **41**, 663–671.
- Steuber-Buchberger, P., W. Wurst and R. Kühn. (2008). Simultaneous Cre-mediated conditional knockdown of two genes in mice. *Genesis* **46**, 144–151.

- van der Krol, A.R., L.A. Mur, M. Beld, J.N. Mol and A.R. Stuitje. (1990). Flavonoid genes in *Petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* **2**, 291–299.
- Vasudevan, S., Y. Tong and J.A. Steitz. (2007). Switching from repression to activation: microRNAs can upregulate translation. *Science* **318**, 1931–1934.
- Weil, D., L. Garçon, M. Harper, D. Duménil, F. Dautry and M. Kress. (2002). Targeting the kinesin Eg5 to monitor siRNA transfection in mammalian cells. *BioTechniques* **33**, 1244–1248.
- Wessenegeger, M., S. Heimes, L. Riedel and H.L. Sanger. (1994). RNA-direct *de novo* methylation of genomic sequences in plants. *Cell* **76**, 567–576.
- Yu, J.-Y., S.L. DeRuiter and D.L. Turner. (2002). RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci.* **99**(9), 6047–6052.
- Zamecnik, P.C. and M.L. Stephensen. (1978). Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci.* **75**, 280–284.
- Zamore, P.D., T. Tuschl, P.A. Sharp and D.P. Bartel. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33.

24 Genomes, Transcriptomes, Proteomes, and Bioinformatics

Rationale	561	Proteomes and proteomics	566
Essential nomenclature	563	Bioinformatics	571
Genomes and genomics	564	Search for genes – have a BLAST!	571
Transcriptomes and transcriptomics	565	References	573

Rationale

The completion of the human genome project generated massive amounts of data that have provided hints as to the structure and organization of the genetic blueprint of the cell. Among the benefits of this monumental undertaking was a first-time-ever ability to associate genome anatomy with genome physiology, i.e., an ultrastructure perspective of the function(s) of the genome, as well as an ability to perform sequence-wise and structure-wise comparisons with other organisms. In the past, positional cloning (i.e., identification of a gene by finding its position on a genetic map), was one of the few imprecise tools available for zeroing in on genes that, when mutated, were responsible for the onset and progression of various diseases. People at risk for various genetic disorders were often identified by the presence of a restriction fragment length polymorphism (RFLP), meaning that DNA from normal and carrier or afflicted individuals would be cut differently by one or more restriction enzymes; the major shortcoming of this method is that not all genetic diseases have an associate RFLP. Thanks to automated PCR-based methods, entire genomes have been sequenced from bacteria, viruses, yeast, plants, and animals, with new organisms being added continually to the ranks. Thus, scientific community regards itself squarely in the postgenomics era, in which elucidation of genome function is the next sought after prize.

It is not surprising, then, that disciplines concerned with the function of the genome (i.e., the RNA and proteins encoded therein) are of immense, if not frenzied, interest. New emergent disciplines and sub-disciplines require new names. It has become rather stylish to simply add the suffix “-omics” to a noun

that has something to do with the flow of genetic information, or analysis thereof, to produce a name for the discipline. Although one could reasonably guess the subject matter associated with the new disciplines, a short working definition for several of these terms is provided in Table 24.1 for the convenience of the reader.

Genomics, transcriptomics, and proteomics are intimately linked, obviously, as professed by the central dogma of molecular biology (reviewed in Chapter 1). It is important to keep in mind, however, that DNA sequence information is only part of the story, because not all genomic DNA has the capability to be “functional”. Recall also that RNA, although a commonly examined parameter of gene expression, does not tell the whole story either because of the enormous turn-over of hnRNA in the nucleus and the fact that mRNAs in the cytoplasm are not guaranteed to be translated. Because genomes are stable but the RNA- and protein complement of a cell are not, it makes perfect sense to study both RNA and protein expression data in order to gain a more complete picture of physiology of the cell under local conditions.

Bioinformatics, a unique blend of basic biology, genetics, and information sciences, is somewhat distinct among the various “-omics” disciplines because it refers to the use of digital technology to analyze DNA sequence data, RNA (and cDNA) sequence data, and protein sequence information. With the advent of supercomputing, *in silico* proteomics is expected to become increasingly prevalent as a tool for predicting the folding and function of novel therapeutic peptides. Inasmuch as the amount of new nucleic acid and protein sequence information continues to be deposited into various databases worldwide, informatics platforms need to interface molecular and clinical data, especially in the context of personalized medicine. It is no exaggeration to state that the drug discovery process has been accelerated by the revolution in information technology, and taking full advantage of these informatics tools is a boon to all areas of research in the life sciences.

Table 24.1 The “-omics”, Showing the Relationships Among the Various Functional Components of Gene Expression

Study of the	is known as	and involves the	and which is
Genome	Genomics	DNA content of a cell	Static*
Transcriptome	Transcriptomics	RNA content of a cell	Changes constantly
Proteome	Proteomics	Protein content of a cell	Changes constantly
Glycome	Glycomics	Carbohydrate content of a cell	Changes constantly
Metabolome	Metabolomics	Overall chemical composition of a cell	Changes constantly

Well-orchestrated expression of DNA-encoded information at all levels results in cells, tissues, and ultimately organisms that have specific, recognizable characteristic

*Mutations and telomere length notwithstanding.

Essential nomenclature

Accession Number: A numeric identifier of a nucleic acid sequence (genomic DNA or cDNA) or an amino acid sequence (protein). The accession number acts as an electronic card catalogue of sorts, facilitating expedient recovery from a database.

Bioinformatics: The marriage of computers, information technology, and biotechnology that allows users to sort through and archive enormous quantities of nucleic acid and protein sequence information. This also extends into the realm of image informatics¹, where data extracted from complex gel- and microarray images are quantified, interpreted, and archived.

BLAST[®]: Basic local alignment search tool software. Perhaps the best known and more widely used bioinformatics software.

BLASTn: nucleotide-nucleotide search for a sequence match. This is the most straightforward type of search.

Contig: a group of clones representing overlapping regions of a genome (contiguous pieces).

Expressed Sequence Tag (EST): A unique segment of cDNA with a base sequence identical to at least part of the coding region of a gene, generally used as landmark for mapping. Succinctly, an EST is a partial cDNA sequence, 200–500 nt, which usually represents the 3' end of the cDNA.

Genome: The entirety of chromosomal DNA found in a cell. In some applications, it may be useful to distinguish nuclear genomic DNA from organellar genomes (e.g., mitochondrial and chloroplast chromosomes).

Genomics: The study of the structure and organization of genes. Genomics is intimately linked to data derived from DNA sequencing.

Chemical Genomics: Chemical manipulation of a genome with organic compounds in order to study the consequences for the cell when cellular pathways are disturbed.

Functional Genomics: The study of RNA levels in the cell and how these levels are modulated as the cell responds to change. Succinctly, it is the study of gene expression.

Metabolome: The overall chemical composition, that is, all metabolites, in a cell or other biological sample of interest.

Metabolomics: Compound profiling; the study of the effects on metabolite levels wrought by the proteins. This type of investigation often suggests possible gene function(s).

Proteome: The full complement of proteins produced by a cell at a particular time; the entire collection of translation products as purified from whole cell extracts.

Proteomics: The study of the protein composition of a cell under certain biochemical conditions. This includes an evaluation of the structure, function, and distribution of expressed proteins and how the assortment of proteins changes in response to natural and artificial stimuli. Proteomics constitutes a systematic means for correlating protein function and abundance with genome function in normal cells as well as the disease state, facilitates characterization of many proteins simultaneously, and has emerged as a major platform for drug discovery.

Transcriptome: The complete set of mRNA molecules produced by a particular cell under a particular set of conditions.

¹IQBase (Media Cybernetics; www.mediacy.com) is an outstanding software solution for this and other applications.

Transcriptomics: The all-inclusive analysis of an organism's transcripts and how the expression of these mRNA molecules changes as conditions inside the cell change.

Genomes and genomics

Sequencing a single gene produces a lot of information. The sequencing of several genes that are of experimental interest to a laboratory produces an even greater amount of information. Sequencing an entire genome produces so much information that it could not possibly be organized, analyzed, archived, or made readily accessible to other investigators without very powerful, high-speed computers, outstanding software for the recovery of desired information from the database, and users who have a good understanding of the science which the database entries represent. One such function of bioinformatics is known as "data mining", which is a unique blend of information sciences and biology. Students who are proficient in these areas can expect exciting careers in diverse areas over the next several years. For example, a major in life science and a minor in information technology, or vice versa, is a good start.

Among the goals of genomics is elucidation of the precise stretches of DNA that constitute discrete genes. One of the primary methods used to accomplish this goal is to examine sequencing data for an open reading frame (ORF) which is a long run of nucleotides that is *not* punctuated by a nonsense (stop) codon. Bonafide ORFs also manifest the requisite start codon (ATG; less frequently GTG), and eventually one of the three stop codons TAA, TAG, or TGA. In eukaryotic cells, an ORF may be punctuated internally with coding- and non-coding domains (exons and introns, respectively), which are easily recognized by the presence of the highly conserved 5' (GT) and 3' (AG) splice junctions and the sequences immediately surrounding the splice sites. Recall that interrupted genes, that is, those consisting of exon and intron sequences, are found in eukaryotic cells, and not in prokaryotic cells, the latter of which contains an absolute economy of genetic information. Certain phages (viruses) such as Φ X174² even exhibit two genes occupying the same stretch of DNA, accomplished by utilizing two different overlapping reading frames. This is not an entirely unexpected finding because (1) the genome of this virus is quite small (5375 nucleotides) and must produce all of the proteins necessary to direct the synthesis of progeny virions; and (2) viruses in general have gone through astronomical numbers of replications resulting in genomes that are sleek and maximally efficient.

For larger-scale sequencing projects the classical phagemid and cosmid vectors are woefully inadequate because the insert sizes that these vectors can accommodate are no greater than about 50,000 base pairs. Efficient mapping of complex genomes is possible using more specialized vectors such as yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs),

²As an historical note: the phage Φ X174 was the first genome ever to be sequenced (1977).

which can accommodate hundreds of thousands of nucleotides. The use of BACs and YACs has allowed sequencing of entire genes, including the flanking promoter and regulatory elements, and has facilitated identification of genetic linkage relationships between loci. These methods are within the realm of structural genomics because the emphasis on genome structure and organization. In contrast, functional genomics is all about the patterns of expression of the genes resident in a cell.

Transcriptomes and transcriptomics

The full complement of mRNAs that a cell is producing is known as its transcriptome. In the broadest sense, this entire book is about the characterization of the transcriptome because of the numerous methods presented for the isolation and characterization of RNA. For example, microarray analysis (Chapter 21) is a contemporary method for assessing which genes are expressed under a defined set of experimental conditions and how expression changes in response to experimental challenge. In a way, microarray analysis might be thought of as a glorified dot-blot analysis – known DNA sequences immobilized on a solid support hybridize cDNA that represents the RNA being produced by the cell at the moment of lysis.

cDNA sequencing data mirror transcript complexity and abundance within a sample. One source of expression data is expressed sequence tags (ESTs), which are partial cDNA sequences. EST analysis can be used to profile mRNA expression in cell culture, in tissue samples, in the form of cDNA libraries (subtracted or otherwise), or any biological source (Adams *et al.*, 1991). Another method, developed in the mid-1990s called *serial analysis of gene expression* (SAGE) (Velculescu *et al.*, 1995), involves the analysis of a large population of very short (9–10 bp) cDNA molecules, produced by cutting much larger cDNA molecules with *Nla* III, ligating adapters, and then cutting the cDNA with *BSM* FI (Fig. 24.1). The resulting modified cDNAs are referred to as tags, each of which represents a defined position within a unique transcript. The tags are ligated together into concatemers, which can be readily sequenced. The frequency with which a specific tag is identified is related to the abundance of the corresponding mRNA. SAGE analysis is a high-throughput method for analysis of gene expression which has an experimental component and a bioinformatics component. Collectively, SAGE and EST³ analyses have produced a vast body of knowledge pertaining to gene expression in many different organisms, leading to the development of very sophisticated databases. One thing that has become abundantly clear from these various forms of expression profiling is that many genes are capable of producing more than one transcript as a result of alternative splicing (splice variants) and variation in the choice of

³By comparison with SAGE, traditional EST analysis requires sequencing of significantly longer cDNAs.

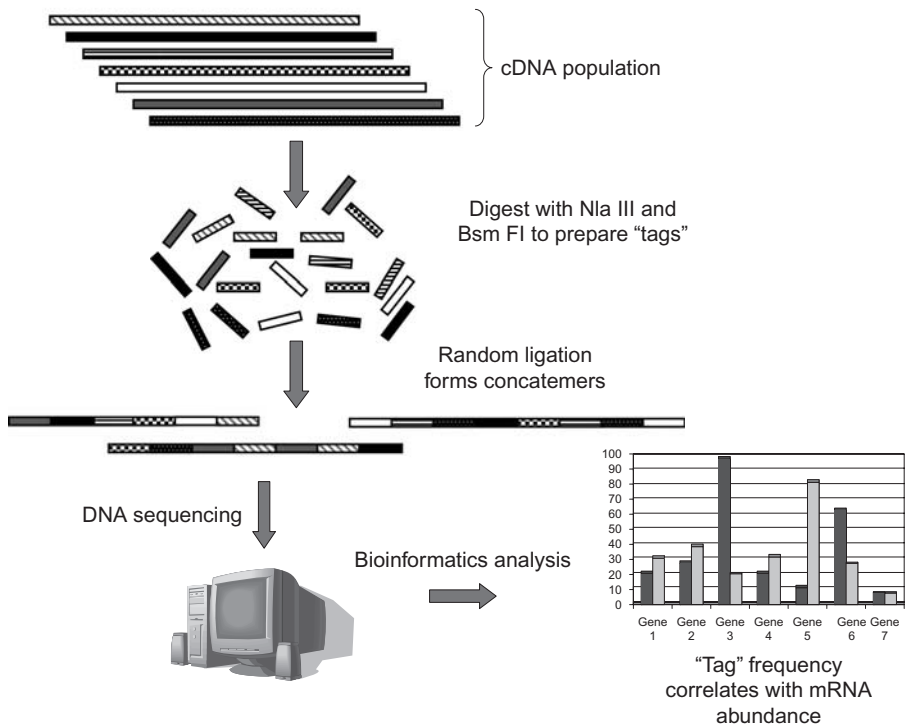


Figure 24.1 Serial analysis of gene expression. The sequential actions of cutting cDNA, ligating adapters, cutting the adapter-ligated cDNA, and randomly ligating the resulting fragments produces concatemers that are sequenced. The frequency with which each "tag" appears is in direct proportion to the frequency of an mRNA species in the original biological material. SAGE is a method for generating a lot of sequencing data very quickly.

transcription start site (TSS). Each of the various transcripts in turn can produce a different protein isoform.

Proteomes and proteomics

Proteins are immensely complex relative to the DNA and RNA sequences that encode them. It has been firmly established that the 25,000 or so resident genes in the human genome encode 1,000,000+ proteins due to variant mRNA transcription start sites, alternative splicing and, especially, posttranslational processing of nascent polypeptides. Posttranslational processing refers to the extensive covalent modifications and related events commonly associated with natural protein maturation. Posttranslational events include glycosylation, acetylation, isoprenylation, phosphorylation, methylation, sulfation, proteolytic

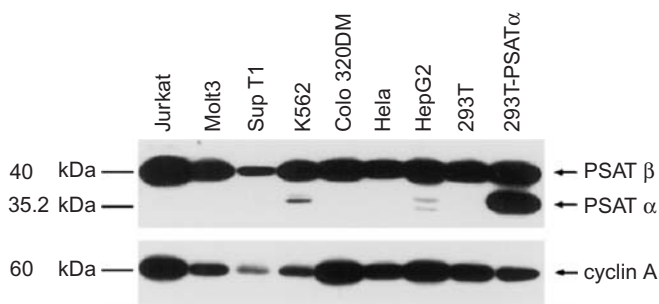


Figure 24.2 Western analysis involves immunological detection of a protein after electrophoresis and blotting. Expression of phosphoserine aminotransferase (PSAT), PSAT, and cyclin A (control) in various human tumor cell lines. The intensity of the band in each lane is indicative of the abundance of the protein in the sample. Reproduced with permission, from Baek, J.Y. *et al.*, (2003). *Biochem. J.* 373, 191-200. © the Biochemical Society (<http://www.biochemj.org>).

processing, complex protein–protein interactions, sumoylation⁴, and ubiquitination, to name but a few. Ergo, the dynamic nature of the proteome is an added level of complexity for proteome profiling in eukaryotic cells. Therefore, proteomics as a discipline has an important role in discerning gene function, and has evolved into the global study of patterns of protein expression in cells and tissues, rather than the analysis of one protein at a time. Protein informatics has become a specialized, formidable subdivision of bioinformatics. Proteomics is important in its own right because amino acid sequence can be used to predict folding, and folding can be used to predict function.

The cell type (prokaryotic, plant, animal, yeast) from which proteins are to be extracted will determine the form of cellular disruption at the onset. Methods include mechanical lysis, enzymatic lysis, homogenization, detergent permeabilization, and grinding. All of these techniques have been described extensively elsewhere, and instructions regarding the precise method to use accompany the purchase of each protein isolation reagent or kit from the various biotech vendors that specialize in protein isolation and characterization.

Methods for protein identification include older methods, newer methods, and older methods that are now enjoying a resurgence. First and foremost, proteins may be characterized by western analysis (Burnette, 1981). Succinctly, proteins are electrophoresed through polyacrylamide gels, electroblotted onto nitrocellulose or PVDF membranes, and then incubated with antibodies to determine the size and abundance of a particular protein in a sample (Fig. 24.2). This method is most often to assay one protein at a time. Two-dimensional (2-D) gel electrophoresis, first described in 1975 (Klose, 1975; O’Farrell, 1975) is one of the premier tools for taking a first look at the proteome. In this method, proteins are first separated by isoelectric point (pI; the pH at which

⁴SUMO (small ubiquitin-like modifier) proteins transiently attach to other proteins and, in so doing, change their function.

the protein has no net charge), followed by electrophoresis perpendicular to the direction of separation of the first dimension; in the second dimension, proteins are further chromatographed based on size. The 2-D gel method permits the simultaneous separation of thousands of proteins, each of which appears as a unique spot upon visualization (Fig. 24.3). A 2-D gel is a fingerprint of the sample; the size of each spot is proportional to the abundance of the protein. Like RNA, proteins are present in the cell at varying abundance levels; therefore, it is important to ensure that superabundant proteins do not mask the presence of low abundance proteins that may resolve in proximity under 2-D electrophoresis conditions. A number of strategies have been devised to fractionate the proteins in a whole-cell lysate in order to maximize resolution of the proteome (Corthals *et al.*, 2000; Zuo and Speicher, 2000). These methods, which are in widespread use, are based upon an orderly fractionation scheme so as to simplify analysis of the sample. Specialized software for 2-D gel image analysis is readily available for automating the identification and quantification of unique spots on the gel, that is, image informatics, and provides an electronic means for archiving the image.

Next, unique proteins are recovered from individual spots on the 2-D gel. These proteins are digested with proteolytic enzymes, rendering much smaller peptides that are identified by mass spectrometry, a popular form of which is

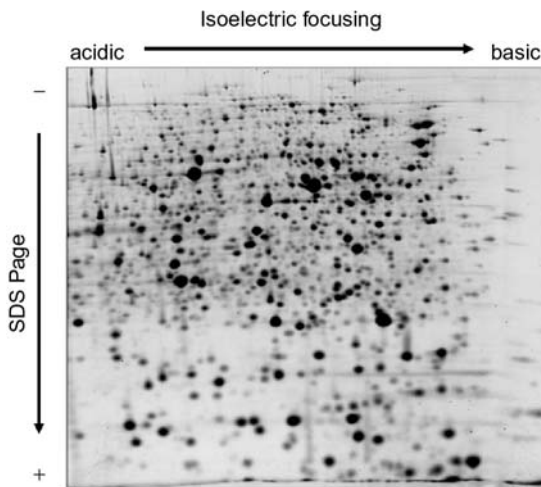


Figure 24.3 Traditional two-dimensional protein electrophoresis. Proteins are first separated by isoelectric point and then, perpendicular to direction of the first separation, chromatographed by mass. The resulting distribution of “spots” is a proteomic finger print of the sample. Much like western analysis, the intensity of each spot correlates with the abundance of the protein at that location. Proteins can be cut out of the gel for identification. Gel image courtesy of Sjouke Hoving and Hans Voshol, Novartis Institutes for BioMedical Research. Generated as part of the Fixing Proteomics Campaign (www.fixingproteomics.org).

known as matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) analysis. Developed in the 1980s, “MALDI analysis” has become an important tool in the area of proteomics. Because it is the DNA that encodes any and all possible proteins, MALDI-derived amino acid sequence info can be used to match the protein under investigation with the nucleotide sequence associated with a specific genetic locus. However, physically mapping a protein to a gene does not simultaneously identify its function. Elucidation of protein function is often a difficult task because a particular protein may function in a collaborative manner with other proteins or may be a component of much larger complex.

At present there are a number of strategies being employed to discover the function of specific proteins, including the yeast two-hybrid system, various permutations on immunoaffinity capture, nuclear magnetic resonance (NMR), crystallization, capillary electrophoresis, high performance liquid chromatography (HPLC), and protein microarrays that are used to detect protein–protein, protein–lipid, and protein–nucleic acid interactions. Much of the data already generated in this area can be accessed through the protein interaction databases; for example, the Molecular Interaction Database (MINT; <http://mint.bio.uniroma2.it>) and the Biomolecular Interaction Network (BIND; www.bind.ca) are two noteworthy examples. Antibody microarrays are now being manufactured, consisting of hundreds of monoclonal antibodies representing a spectrum of biological functions, including “themed” arrays for various infectious diseases, cancer, and apoptosis. One popular format for the detection of protein expression involves labeling proteins from a biological source with Cy3 and Cy5 dyes, and then hybridizing these labeled proteins to the arrays which contain the antibodies. This dual labeling scheme can also be applied to the 2D-gel format described above (Fig. 24.4). For example, proteins in one sample can be labeled with Cy3 and proteins in another sample (control vs. experimental) can be labeled with Cy5. The two protein population are mixed together and run on a single 2D gel. Laser excitation of Cy3 will reveal “green” labeled proteins while laser excitation of Cy5 will reveal location of “red” labeled proteins. One of the nice things about digital image analysis is the ability to not only overlay images but also to subtract what two or more images have in common. Thus, image subtraction (common locations of red and green fluorescence) will immediately reveal the location(s) of proteins uniquely expressed in each population. This method precludes the requirement for comparing X-ray films from two different gels side-by-side to find differences.

Other more traditional systems for proteomics investigations are various transcription-translation systems, including wheat germ extract, rabbit reticulocyte lysate, canine pancreatic microsomal membranes, and proteases, to name a few.

It would seem that the number of proteomics databases is increasing rapidly, although unlike GenBank and other databases in the public domain, access to some privately controlled proteomics databases requires licensing and/or access fees which purportedly are used to ensure maintenance of the databases as well as future development. The professional organizations related to proteomics, which are good starting points for novices in the discipline, are the Proteome

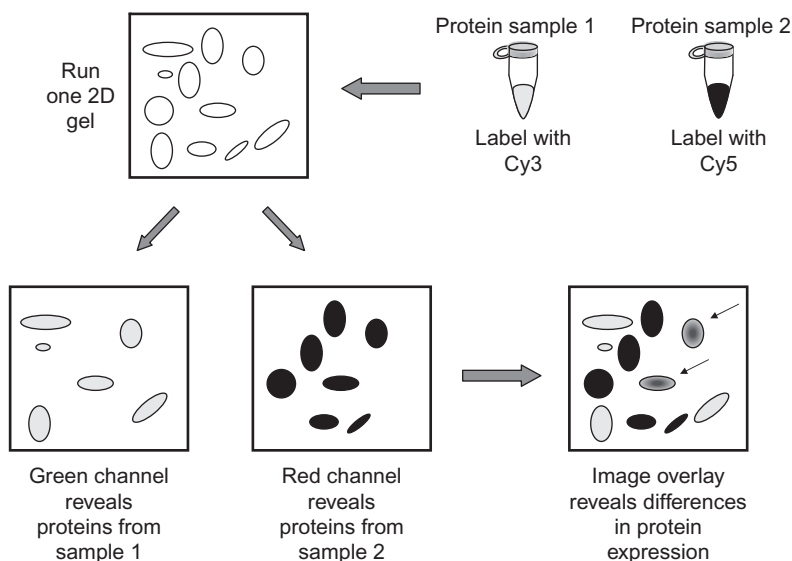


Figure 24.4 Fluorescence two-dimensional protein analysis. Two protein samples individually labeled with Cy3 and Cy5 and electrophoresed on one gel. Laser excitation then reveals the locations of Cy3- and Cy5-labeled proteins. By digitally overlaying each image and subtracting the areas of common fluorescence, differences in protein expression among the samples are revealed.

Society (www.proteome.org), the Microarray Gene Expression Data Society (www.mged.org), and the Human Proteome Organization (www.hupo.org). In addition, the variegated nature of posttranslational modifications of proteins, and methods for the study thereof, has been recently highlighted (Packer *et al.*, 2008; Witze *et al.*, 2007; Mann and Jensen, 2003).

It is worth mentioning in this brief discussion of proteomics that peptides, like DNA and RNA oligonucleotides, can be made-to-order using amino acids with various posttranslational modifications. Applications involving the use of customized peptides include epitope mapping for immunology applications, identification of active peptides for new drugs, identification of enzyme substrates and inhibitors, characterization of protein–protein interactions, characterization of protein–nucleic acid interactions, and development of novel antibodies for vaccine development.

Metabolomics, essentially compound profiling in the cell, in the blood, or in another fluid, is the logical extension of proteomics. Unlike proteomics, metabolomics is concerned with the changes that are wrought by the various proteins produced by the cell. These changes correlate directly with the precise posttranslational modification(s) of the protein and the local environment. The field of metabolomics originated more than three decades ago, and is now enjoying something of a resurgence because of the large amount of protein sequencing information being added to protein databases daily. Metabolomics

Table 24.2 Important Bioinformatics Web Sites

NCBI	www.ncbi.nlm.nih.gov	GenBank
EMBL	www.embl.org	EBI core databases
Ensembl	www.ensembl.org	Vertebrate genomes
TAIR	www.arabidopsis.org	Arabidopsis database
DDBJ	www.ddbj.nig.ac.jp	DNA Data Bank of Japan
UniProt	www.uniprot.org	Protein sequence database
ArrayExpress	http://www.ebi.ac.uk	Microarray database
Kazusa	http://www.kazusa.or.jp/codon/	Codon utilization tables

is useful for evaluating a drug's toxicity, potential side effects, and overall safety. Methods utilized to profile the metabolome include NMR, mass spectrometry, and various other traditional chromatography methods, including liquid and gas chromatography.

Bioinformatics

Vast new amounts of novel nucleic acid- and protein sequencing data are produced daily, most of which eventually becomes part of the public domain. As such, the National Center for Biotechnology Information (NCBI; Bethesda, Maryland) was established for the expressed purpose of making these data available to anyone wishing to access it, and at the same time provide the proper software tools to support queries into the database. Similarly, the European Molecular Biology Laboratory (EMBL) (Heidelberg, Germany), and the DNA Data Bank of Japan (DDBJ) (Mishima, Japan) coordinate their efforts with NCBI to provide ready access of this massive quantity of data⁵ to investigators worldwide. A variety of other specialized databases are readily accessible from these central portals and Table 24.2 lists some important bioinformatics web sites. Beginning with the NCBI homepage, one may search for specific DNA sequences (Entrez), search for the chromosomal location of genes of interest (Human Map Viewer), and search through scientific literature related to a specific gene (PubMed). A number of other tools are available at these sites, with which the reader should become familiar.

Search for genes – have a BLAST!

The ever-increasing wealth of information resulting from genome, cDNA, and protein sequencing represents a daunting challenge to investigators who wish

⁵As of February, 2009 GenBank contains approximately 101.5 billion base pairs of data representing nearly 102 million different sequences. Check out the statistics anytime at www.ncbi.nlm.nih.gov/genbank/genbankstats.html

to utilize this information: how does one efficiently search and recover specific items from databases consisting of billions of bases? To do this, various computer programs have been written that simplify the tasks of searching for, and identification of, DNA or protein sequences generated by the investigator. Perhaps best known is Basic Local Alignment Search Tool software, better known as BLAST[®], a powerful and versatile tool that searches databases for nucleic acid or amino acid sequences that match the input sequence and simultaneously shows sequence alignment. Another favorite is Gene Recognition and Assembly Internet Link, also known as GRAIL (<http://compbio.ornl.gov/grailexp>), software that identifies entire genes, exons, promoters and promoter elements, and other features within DNA sequences provided by the user.

The BLAST program can be accessed directly at www.blast.ncbi.nlm.nih.gov. Many other internet sites have links to BLAST. The process for performing a BLAST search is simple, and the most straightforward method is to select “nucleotide-nucleotide BLAST” (blastn). Users are asked to either manually enter or cut-and-paste the desired search string (the nucleotide sequence being investigated). Because of possible confusion due the similar appearance of the uppercase letters “C” and “G”, the convention now is to enter a base sequence as all lowercase letters, a, c, g, t, in the search box. The user then has the option of searching the entire unknown sequence or portions thereof by entering the first and last nucleotides in the “FROM” and “TO” windows, respectively. The user is also asked to select a database to be searched. Unless there is a compelling reason to pick a specific database it is usually best for relatively new users to select the default non-redundant database “nr” which will perform the most comprehensive search possible⁶. Finally, click on the BLAST icon to begin the search process. Depending on a number of variables, the search may require anywhere from seconds to minutes, or longer, to complete. The BLAST software will provide an indication of the approximate time needed to complete the search and will also provide a request ID number, which is a reference number for the transaction. When the search has been completed, another window will then open another window in which the results of the BLAST search will be presented. When the search results are finally available, there will be two scores assigned to each sequence: (1) the *S score* (bit score), a statistical value associated with the number of matches between the sequence provided by the user and the match(es) retrieved from the data base; and (2) the *E value* (expectation value), which is the number of matches with the same *S score* that one would expect to find by chance. Thus, for any BLAST search, the higher the *S score* and the lower the *E value* (should be much lower than 1.0), the better the match. In the event that a DNA sequence is not available for performing a BLAST search, the NCBI website has other extensive, very impressive resources with which one can search for the genes that have been implicated in the onset and progression of various diseases, and will even provide a detailed genetic map.

⁶EST and other specialized databases are also available for some organisms.

Table 24.3 Select Amino Acid Folding Propensities

	α -helix formation	β -sheet formation
Favorable tendency	Alanine Arginine Glutamic acid Leucine Methionine	Isoleucine Phenylalanine Tryptophan Tyrosine Valine
Unfavorable tendency	Asparagine Glycine Proline	Aspartic acid Asparagine Glycine Glutamic acid Proline

Examples of certain amino acids are associated with specific types of intramolecular protein folding. These classifications are generalities only; the role of amino acids in 2° protein structure varies greatly depending on the local conditions, including side-chain interactions. Compiled from data published by Pace and Scholtz (1998), Farzad *et al.* (2008), and Zvelebil and Baum (2008).

In order to become familiar with the proper and expedient use of the BLAST program, one can go to any of the above-referenced web sites at any time and practice using the various information tools. A convenient way to approach a first-time practice session is to enter either a known sequence, perhaps related to a gene being studied in the lab, or the PCR primers being used to amplify portions of the gene. Alternatively, one can simply make up a sequence for the purpose of gaining familiarity with the search process. Also noteworthy is the relatively new Primer-Blast tool which can be used to design your own primers for your own template using your own parameters. Give it a try at www.ncbi.nlm.nih.gov/tools/primer-blast.

Bioinformatics tools are also used to predict the folding of proteins, based on the 1° structure (the order of amino acids), from which it may be possible to infer function. For example, the 2° level of protein folding involves α -helices and β -pleated sheets, and certain amino acids tend to favor the formation of one of these 2° structures over others (Table 24.3); consequently, the path of the assembled amino acid determines the 3° structure that protein. There are also tools for the examination of protein three-dimensional (3-D) structure based on X-ray crystallography data, and a great many of these resources are available to anyone free-of-charge. It is enlightening to visit the NCBI website and explore the possibilities!

References

- Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, *et al.* (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651–1656.

- Burnette, W.N. (1981). "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analyt. Biochem.* **112**, 195–203.
- Corthals, G.L., V.C. Wasinger, D.F. Hochstrasser and J.C. Sanchez. (2000). The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **21**, 1104–1115.
- Farzad, F., N. Gharaei, H. Pezeshk and S.A. Marashi. (2008). β -Sheet capping: signals that initiate and terminate β -sheet formation. *J. Struct. Biol.* **161**, 101–110.
- Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26**, 231–243.
- Mann, M. and O.N. Jensen. (2003). Proteomic analysis of post-translational modifications. *Nature Biotechnol.* **21**, 255–261.
- O'Farrell, P.H. (1975). High-resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021.
- Pace, C.N. and J.M. Scholtz. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75**, 422–427.
- Packer, N.H., A.A. Gooley and M.R. Wilkens. (2008). One gene, many proteins. In *Proteome Research: Concepts, Technology, and Applications* (M.R. Wilkens, R.D. Appel, K.L. Williams and D.F. Hochstrasser, Eds). Springer Publishers, New York. pp. 95–122.
- Velculescu, V.E., L. Zhang, B. Vogelstein and K.W. Kinzler. (1995). Serial analysis of gene expression. *Science* **270**, 484–487.
- Witze, E.S., W.M. Old, K.A. Resing and N.G. Ahn. (2007). Mapping protein post-translational modifications with mass spectrometry. *Nature Methods* **4**, 798–806.
- Zuo, X. and D.W. Speicher. (2000). A method of global analysis of complex proteomes using sample prefractionation by solution isoelectrofocusing prior to two-dimensional electrophoresis. *Analyt. Biochem.* **284**, 266–278.
- Zvelebil, M. and J.O. Baum. (2008). Obtaining secondary structure from sequence. In *Understanding Bioinformatics*. Garland Science, New York. pp. 413–462.

25 An RNA Paradigm

A typical experiment? 577
Sensitivity issues 580

What to do next 580
Where to turn for help 582

The protocols contained in the preceding chapters of this text illustrate various methods for the expedient isolation of RNA from a variety of biological sources and its subsequent analysis. The judicious application of various combinations of these techniques constitutes a systematic method for understanding the role of transcriptional products in the regulation of gene expression. Any experimental model can be well-characterized by assaying specific mRNAs as one parameter of gene expression, although having protein expression data from Western analysis, immunochemical assay, or proteomics array analysis, generally provides a more complete understanding of gene regulation under the prevailing conditions. A typical RNA-based experimental design includes an assessment of steady-state levels of RNA and, perhaps, the rate of transcription in control and experimental cell populations. These techniques are performed to assign observed changes in the prevalence of certain mRNAs to the transcriptional level or due to some posttranscriptional event(s).

Northern analysis of cytoplasmic RNA affords the investigator an opportunity to directly examine specific members of a transcript population, albeit one or a few mRNA species at a time. Classical Northern analysis is limited, in part, by the very mechanics of the assay. It is well established that the physical immobilization of target RNA on a solid support (e.g., a nylon membrane) in some way prevents complete accessibility of those molecules to a nucleic acid probe. However, it is Northern analysis alone that reveals the overall length of the mRNA, as produced by the cell. All of the other methods for RNA analysis, no matter how high-tech sounding, analyze only a portion of the full-length transcript. Moreover, through the independent study of nuclear RNA and cytoplasmic RNA, very definitive conclusions can be drawn about the transcriptional and posttranscriptional regulation of gene expression.

A more complete format for transcript quantification is known as solution hybridization, epitomized in the form of the nuclease protection assays and, of course, the polymerase chain reaction. The enhanced sensitivity of these techniques commonly precludes transcript enrichment in the form of poly(A)⁺ selection, which could even be counterproductive. The results of these experiments

are often quantified by autoradiography, chemiluminescence, fluorescence, or scintillation counting, coupled with digital image analysis. RNA characterization is most often associated with some form of assessment of gene expression, so greater sensitivity and resolution are always desirable. RT-PCR and array analysis, and innovations in cDNA sequencing have facilitated profiling the depths of the cellular biochemistry to levels inconceivable 15 years ago and continue to offer greater potential in the form of the many permutations of this technique.

The induction or repression of transcription is commonly reported in the literature in terms of *relative abundance*, a term used to compare the observed prevalence of specific mRNA species in control and experimental cell populations. For example, it might be reported that contact inhibition results in a 50-fold decrease in the abundance of *c-myc* oncogene mRNA in diploid fibroblasts, compared to the abundance of *c-myc* mRNAs in exponentially proliferating, subconfluent cultures of the same. Thus, statements can be made about the increase or decrease in the prevalence of transcripts without having knowledge of the absolute mass of these transcripts in the cell. If desired, however, the starting mass of a sample can be accurately gauged by performing real-time PCR, in which product mass is assayed as it accumulates in the reaction tube and is related to a known amount of starting material through the use of a standard curve.

Molecular scrutiny of a model system does not necessarily stop at transcriptional characterization. Having assayed the prevalence of transcriptional products and, perhaps, the relative transcription rates of different genes by the nuclear runoff assay, an investigator may wish to begin to characterize the molecular basis of observed variations from control or base-line expression. Gene expression is, after all, only partly about the biogenesis of RNA. At the level of genome organization, the modulation of mRNA may be, in part, the result of a gene rearrangement. It is entirely possible that an aberration within the coding portion of a locus, or the flanking sequences that influence its expression, constitute at least one underlying basis for the up- or down-regulation observed upon transcript quantification. Subsequent investigations may involve Southern analysis to probe for such structural changes or DNA sequencing to look for point mutations, the latter of which has been greatly enhanced by the field of bioinformatics. This is possible because of the completed sequencing of the human- and other genomes and the fact that immense numbers of organism-specific sequences may be screened using microarrays.

Gene expression at the cellular level culminates in the form of a functional polypeptide. To provide clearer definition to an experimental system, characterization of translational and posttranslational profiling of gene expression adds yet another dimension to understanding the biochemistry of the cell. Although the presence of one mRNA species or another suggests that they are translated, experimentally-induced challenges may compromise access of these same mRNAs to the translational apparatus; thus Western analysis at some point is in order. Further, with the advent of RNAi, it is now possible to study the effects of transiently blocking protein synthesis by targeting the mRNA template for destruction. One must also be mindful that the posttranslational

modifications commonly associated with eukaryotic gene expression may also be compromised under the parameters defined in a particular experimental situation. All of the new proteomics tools are readily available to help the investigator gain a clearer understanding of the role of protein maturation and, ultimately, gene function. Thus, appraisal of the cellular biochemistry from several perspectives is generally in order.

A typical experiment?

There is no such thing as a typical experiment. Before investing countless hours and thousands of dollars investigating a particular problem with various expensive kits and high-tech approaches, a simple initial characterization of a new system is always in order (Fig. 25.1). The type of “quick and dirty” assessment required is a direct function of the model being proposed. For example, an abundant transcript might easily be assayed from large numbers of samples or across multiple time points by something as unsophisticated as dot blot analysis, though this method is lacking in the qualitative component that

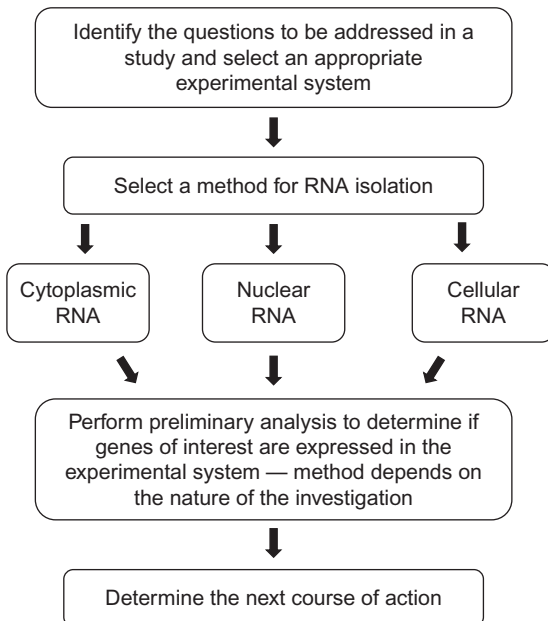


Figure 25.1 Preliminary analysis is intended to determine whether a selected experimental system is appropriate. In-house expertise, equipment, and the nature of the questions being asked determine the most reasonable “quick and dirty” approaches. Perhaps the most important information to be gleaned from preliminary characterization is whether the investigator is capable of isolating RNA of sufficiently high quality from the biological source to support the intended experiments.

associated with the electrophoresis of RNA. If transcripts of interest are not detected, it may be necessary to modify the experimental design. Experimental modifications could include (1) selecting an alternative method of RNA isolation; (2) selecting an alternative biological source; (3) selecting a more sensitive assay; or (4) induction of the genes of interest by chemical or physiological stimulation. At the same time, there is nothing wrong with attempting to generate preliminary data using supersensitive RT-PCR methods; the choice really depends on the nature of the starting material and the questions being asked.

In an initial characterization, analysis of total cellular *and* total cytoplasmic RNA may be very useful, because many transcribed RNAs are degraded in the nucleus (hnRNA) and hence neither mature nor appear in the cytoplasm in the form of mRNA. Further, if a probe hybridizes to cellular RNA samples and not to cytoplasmic RNA samples, this suggests some type of posttranscriptional regulatory mechanism in action. If initial assessment yields encouraging data, a more detailed examination of the system is in order.

The Northern analysis is a time-honored tool with which to examine transcript size (Fig. 25.2), although few investigators use this technique for quantitative purposes any longer because of the *much* greater sensitivity of other

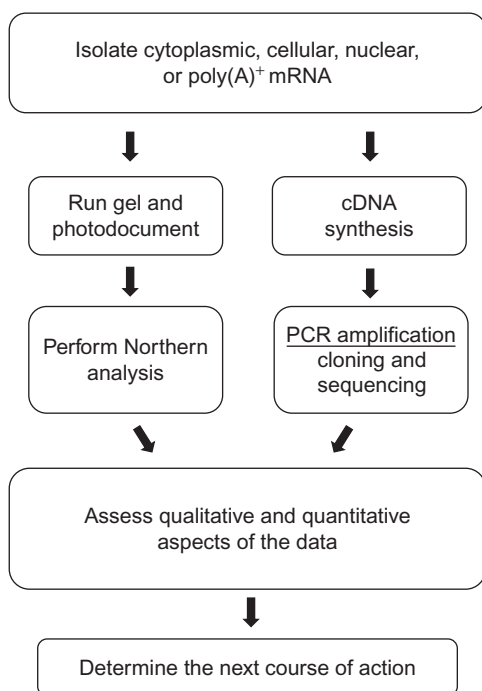


Figure 25.2 Northern analysis can be used to determine native transcript size and look for the presence of multiple transcripts. It is generally advisable to perform a side-by-side analysis of the RNA samples with some type of PCR-based assay. This strategy is invaluable if the abundance of the transcript is below the level of detection by Northern analysis.

available methods. Knowledge of the size of the transcript as well as its abundance demonstrates the specificity of the probe being used and at the same time permits direct comparison to previously published data (if any) on the size and abundance of the messages under investigation. Be advised, however, that a great deal of contemporary research involves the assay of genes that are undetectable by standard blot analysis and RT-PCR may be needed to demonstrate that the RNA is capable of being assayed (is of sufficiently high quantity) in the event that no signal is generated from the blot and, of course, that the RNA prep is of sufficiently high quality.

As suggested above, the more sensitive quantitative assays of transcription are those based on solution hybridization, such as in nuclease protection and the various microarray- and PCR-based assays (Fig. 25.3). It may also be prudent to include one or more experiments for the study of the rate at which various transcripts are produced in control cell populations compared to experimentally manipulated cells. This is readily accomplished by performing the nuclear runoff assay. Transcription rate data are then examined along with

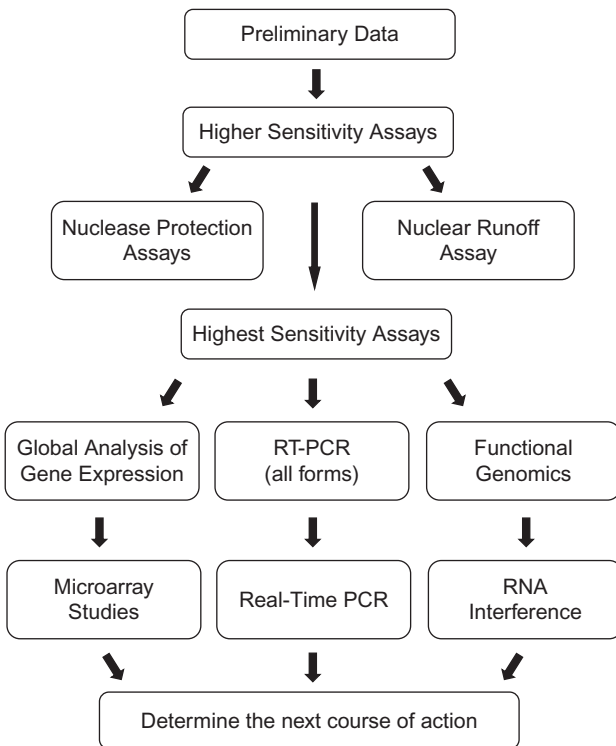


Figure 25.3 The nuclease protection assays and nuclear runoff assay offer greater sensitivity than filter-based assays (e.g., Northern analysis). For maximum experimental latitude, sensitivity, resolution, and complementation of techniques, microarray screening, real-time PCR, and gene suppression by RNA interference are all state-of-the-art methods.

data from steady-state transcript assays, from which conclusions are drawn. It is also possible that perceived changes in gene expression might be the result of a structural change or rearrangement in the locus itself or in the flanking regulatory elements that influence its expression, and that possibility can be investigated by Southern blotting or microarray analysis.

Superior sensitivity and resolution are accolades attributed to PCR. This revolutionary approach to the amplification of cDNA and genomic DNA sequences is among the surest of methods to assay transcription in virtually any experimental context. Of course, PCR requires transcript- or gene-specific sequencing data to generate useful primer sequences and there are numerous strategies described in this book and bioinformatics tools online, which address this task. Although RT-PCR can be a stand-alone technique in the assay of gene expression, it is wise to seek corroborating data using one or more of the previously described, more traditional methods.

Sensitivity issues

This laboratory guide presents a number of methodologies for the isolation and subsequent characterization of RNA. It should be abundantly clear that each of these methods offers distinct advantages and disadvantages which extend into the realm of the overall reliability and sensitivity of the particular assay.

Most often, levels of RNA in an experimentally related set of samples are described in terms of relative abundance, meaning that one of the samples is designated as the control or baseline sample and assigned a value of 1.0 and to which each of the other samples are compared. Be aware, however, that a perceived relative abundance is directly related to the relative sensitivity of the assay at hand, and evaluating the same set of samples using a battery of different methods will often produce startlingly different results.

The relative sensitivity of several common transcription assay methods was described in Chapter 19 as Farrell's RNA Sensitivity Index, in which RNA-based assays are ranked from highest sensitivity (real-time PCR) to lowest (dot blot analysis). Keep in mind that techniques that rank lower on the sensitivity index are nonetheless able to provide useful information. Consider, for example, the classical transcription assay technique of Northern analysis which provides information about the full size of an mRNA, a datum not associated with methods of higher sensitivity ranking.

What to do next

The observation that one or more RNAs of interest have been induced or repressed may be the basis for a decision to expend further laboratory resources to continue to characterize the "phenomenon". Gene expression is by its very nature multifaceted, and the experimental design and associated timeline will,

in great measure, be determined by the particular questions being asked in an investigation. Moreover, unexpected data may suggest a new research direction which may include one or more of the following areas:

- Mapping the transcription start site(s) for a gene under a defined set of experimental conditions.
- Using RNA interference to discover the precise function of a gene.
- Recovering flanking DNA sequences from a genomic library to eventually reconstruct the entire gene as well as the promoter region.
- Determining the exon-intron structure of the loci under investigation.
- Assessing of the splicing pattern of hnRNA which might suggest alternative mRNA transcripts from the same genetic locus.
- Analyzing the gene and its flanking sequences for possible protein binding sites via gel shift assays, DNA footprinting, or proteomics arrays.
- Performing site-directed mutagenesis of either the structural portion of the gene, the regulatory sequences, or both, to discover the consequences of over- and underproduction of a wild-type protein or an improperly folded variant. This can be extended to the organism level as well.
- Performing *in situ* hybridization, including *in situ* PCR, to determine the histological distribution of cells expressing genes of interest.
- Developing PCR primers for diagnostic applications.
- Determining the cellular distribution of corresponding protein products.
- Performing high-level protein expression in prokaryotic systems and eukaryotic systems for clinical or biophysical studies.
- Studying putative regulatory elements by ligation to so-called reporter genes.
- Performing homologous recombination of the gene under investigation with undifferentiated cells in order to examine the role of the locus in differentiation and development.
- *Responsibly* producing transgenic animals to facilitate *in vivo* production of life-saving pharmaceuticals.

If, at some point, it is deemed necessary to retrieve a full-length sequence from a previously existing library, the investigator is cautioned that older libraries (i.e., those constructed between the late 1980s and the mid-1990s) may be severely lacking in longer cloned sequences and are not likely to be as representative as newer libraries. There are two principal reasons for this. First, older libraries were commonly synthesized via oligo(dT)-priming of poly(A)⁺ mRNA. Inefficient synthesis of first-strand cDNA commonly occurs when primed in this manner, resulting in a marked reduction in the size and variety of cDNA molecules; a lack of sequence in the region corresponding to the probe obviously will not produce a hybridization signal. In more extreme cases, many mRNAs might not be represented in the library at all. Second, oligo(dT)-primed libraries are under-represented because poly(A)⁻ mRNA would not be represented in the library. More rationally, the synthesis of first-strand cDNA is primed using oligo(dT) *and* random primers for completeness, as described in Chapter 17. In fact, entire libraries are now made by PCR.

Libraries can be screened by traditional approaches, such as plating out an aliquot of the library, then performing a plaque screening assay using nucleic

acid probes or, if applicable, antibodies. If some sequence information is available, either from studies of a closely related member of the same gene family or from the same gene in another species, it may be possible to synthesize primers that will support screening the library by PCR. Failure to recover specific members of the library by traditional screening approaches or by PCR either mandates the design of a new probe or new primers or may well necessitate construction of a new library. The utility and significance of a cDNA library are derived from the fact that it is a representative permanent biochemical record of the cell at the moment of lysis.

Once fished out of a library or otherwise acquired, a cDNA sequence must be identified by performing DNA sequencing and through the use of the vast bioinformatics tools available online to anyone at any time. Once unambiguously identified, a gene sequence, or portion thereof, can be propagated for use as a hybridization probe for Northern and Southern analysis, nuclease protection, and many other applications. Inasmuch as a great deal of the eukaryotic genome is believed to be transcriptionally silent, cDNA sequencing is a rapid and inexpensive means of identifying discrete loci and measuring genetic expression. When discovered, previously unknown cDNA sequences, or cDNA sequences previously unknown to be associated with a particular phenotype, response, or condition, can be added to GenBank for the benefit of all.

Where to turn for help

In this age of technology, bioinformatics, email, and all types of online tools, it is not difficult to perform a literature search or do a great deal of background reading without leaving one's desk. PubMed is a very quick way to access a great deal of scientific literature. This resource can be accessed indirectly from a number of websites worldwide, by going to the NCBI web site (www.ncbi.nlm.nih.gov) or by simply typing www.pubmed.com into a web browser. Further, most universities and government agencies subscribe electronically to large numbers of journals for easy access.

There are a number of different Internet-based bulletin boards where scientists with difficulties can post specific questions and hope for an answer from the readership. It is very prudent, however, to not believe everything you read, especially on the Internet. Periodic visits to these question-and-answer sites often reveal a disproportionate number of questions answered by the same parties all the time, leaving this Author rather suspicious of individuals who apparently have so much time on their hands.

Most biotech companies have easily accessible technical services departments, all of which can be reached by dialing toll-free numbers and many of which are listed in Appendix P. One should take full advantage of tech services and not hesitate to request replacement materials when a product does not perform in the laboratory as advertised. Moreover, one should not be shy about

approaching a company and requesting a significant discount on the first-time purchase of a new kit or reagent. This will provide the laboratory with the opportunity to try something new at a minimal cost. Clearly, if the kit works well, additional kits and related products will be purchased, and all the vendors know this. One should be polite, but assertive.

One of the nicer things about the Internet is that most biotech vendors have posted protocols and user guides online, mostly in the form of pdf files. Happily, gone are the days of having to wait for user guides and protocols by mail or by fax. When the notion comes to mind that a product may be just what is needed to increase the productivity of the laboratory, it is very easy to print out the user guide without leaving one's desk.

If setting up a new lab, looking to expand the current lab operation, or simply trying to identify which products are available from which vendors, there are two excellent on-line resources. First, one may search for just about anything at *www.biosupplynet.com*. Although there are companies that pay a nominal fee to be listed as "Featured Companies", any vendor that has registered with Biosupplynet will show up if they carry that particular product line. One may search by product, company, protocol, or kit. Second, Linscott's Directory of Immunological and Biological Reagents (*www.linscottsdirectory.com*) is another valuable online resource. Of course, one can always go to a search engine and enter the name of the product of interest, although the two above-referenced sites are both very complete and will save a lot of time and aggravation.

Lastly, there are companies that specialize in biotech education. These companies specialize in laboratory-intensive hands-on training. Most of these workshops are presented by bench-current, PhD-level scientists. Some workshops focus on a particular theme, while other workshops are broad-based in the techniques that are presented. In particular, a number of outstanding regularly scheduled workshops are always listed at *www.DNAtech.com*.

This page intentionally left blank

Epilogue

A Few Pearls of Wisdom

With all its diverse applications, the essential character of molecular biology is the judicious use of fewer than two dozen standard reactions. Upon mastery of these techniques, many of which have been presented here, the research potential is limitless. Molecular biology *is not* about kits; rather, it is about planning and determining how best to use the tools at one's disposal to get the job done. A thoughtful, well-considered strategy and the implementation of the correct techniques, in the correct order, will ultimately govern the outcome of an investigation. In this regard, never forget this Author's definition of the 5 Rs of molecular biology:

Rapid
Representative
Reproducible
RNase-free
Reliable

The following is a list of suggestions, some obvious and others not so obvious, that may help streamline the day-to-day activities in the lab and avoid some of the difficulties that may otherwise be encountered.

1. Get rid of powdered gloves. The powder can inhibit PCR, increase background on filter membranes used for nucleic acid hybridization, and causes an all-around mess.
2. Prepare recipe cards that delineate the exact formulation of commonly used buffers and solutions. Having a card at hand will save a great deal of time when reagents are in short supply or when they must be prepared quickly in the middle of an experiment.
3. Read the label. Just as one would not want to be administered the wrong medication, the use of the wrong reagent can put an immediate end to one's experiment. In fact, read the label twice.
4. Be sure to label tubes unambiguously. Even seasoned investigators often need to be reminded of the importance of being able to associate a particular tube with a particular entry in a lab notebook. Consider, for example, using different color markers to label sample and reaction tubes, and record the color of ink in the notebook. In the case of complicated or multistep protocols, writing the step number on the top of the tube, and other salient information on the side of the tube, is very helpful in terms of organization and tube retrieval from the freezer.
5. All stock and working solutions in the lab should be labeled with a predetermined expiration date, in addition to the date of preparation. Be sure to check the expiration

date of everything before using it! A sad consequence of the use of out-of-date and even worse, ancient solutions, in the lab is the unintentionally introduction of RNase from microorganisms that may have taken up residence in these bottles.

6. In this lab, RNA samples are routinely “cleaned up” by redissolving them in the lysis buffer that accompanies spin column RNA isolation kits. This sample is then processed as if working with a whole cell lysate and the results of this type of cleanup are remarkable. This method is particularly useful for changing the buffer in which the RNA is dissolved and for concentrating the RNA *without* having to precipitate the sample.
7. Running a gel is the single best diagnostic that can be used to assess the integrity of a nucleic acid preparation. A small aliquot of sample electrophoresed on a minigel (e.g., 7 × 8 cm) will reveal the probable usefulness of the sample in downstream applications. This can be performed at several points in a series of experiments to check for degradation of RNA, improper denaturation, and efficient conversion into cDNA.
8. If denaturing RNA with formaldehyde for any application, be aware that the sample will quickly become degraded if the formaldehyde pH is below 4.0. Always deionize a fresh aliquot of formaldehyde (and formamide) just prior to use (see Appendix H).
9. If loading buffer containing bromophenol blue changes from blue to green when added to a sample, there has been a radical pH change in the tube, and the sample integrity itself is at risk.
10. When RNA is to be recovered from a lysate that also contains DNA, organic extraction at pH below 6.0 will drive DNA to the aqueous-organic interface. Keep in mind, however, that samples isolated in this manner are almost certainly tainted with genomic DNA. DNase I treatment is necessary, especially for techniques involving PCR.
11. If denaturing RNA with dimethyl sulfoxide (DMSO) and glyoxal, especially for dot blot analysis, be aware that DMSO dissolves nitrocellulose.
12. Wash nucleic acid pellets with 70% ethanol to remove excess salt. This is necessary because nucleic acid molecules and salt form aggregates in solution that demonstrate dramatically reduced solubility in alcohol. Thus, the precipitate at the bottom of the tube is both salt and nucleic acid. Washing the pellet with 70% ethanol will remove much of the salt; however, it is not sufficiently aqueous to redissolve nucleic acids. After two or three washes, do a final wash with 95% ethanol to facilitate rapid drying of the pellet.
13. RNA pellets do not adhere to the side of microfuge tubes as firmly as DNA pellets. When decanting an alcohol supernatant, keep the pellet in sight at all times. If afraid of losing a pellet when decanting, pour off the supernatant into another microfuge tube so that the pellet will not be lost if it slides out of the tube.
14. When pipetting microliter volumes, it can be helpful to pipette individual reagents onto the inside wall of the microfuge tube, close to, but not at the bottom. The individual reagents can then be mixed by pulse centrifugation. This avoids the problem of inaccurate volume delivery to the tube due to capillary action in the tip of the micropipette, and shows convincingly that small aliquots have actually been delivered into the tube.
15. Always test enzyme and reaction components when opening a new package. Non-functional enzyme translates into one thing: no results!
16. If preparing a master mix for multiple reactions, be sure to multiply the “per reaction” volumes by one more than the total number of required reactions. This is a

standard procedure to ensure an adequate volume of master mix in the event of subtle micropipetting errors. This procedure applies to the preparation of cDNA master mixes (first- and second-strand) and PCR master mixes.

17. Get in the habit of making up master mixes for cDNA- and PCR-related applications. This is *especially* important for quantitative analyses. Even subtle variations in microliter reaction volumes or nanomolar concentrations can have a strong negative impact on the accuracy of the resulting data.
18. Pay attention to the final enzyme concentration. Recall that molecular biology enzymes are often stored in 60% glycerol. The volume of a reaction occupied by the enzyme should not exceed 5% of the total volume. Beyond this amount, the enzymes will not function optimally.
19. Consider the use of REDTaq™ (Sigma) to ensure that (1) the enzyme has been added to the reaction tube; and (2) the enzyme has been properly mixed within the reaction tube.
20. PCR optimization is always easier if a thermal cycler with a heated lid is used. This is true because of the small volumes involved, the elimination of mineral oil, and the potentially large number of samples involved. Moreover, many investigators have reported a marked decrease in amount of product generated when a PCR reaction is overlaid with mineral oil.
21. Nucleic acid sequences are always written 5'→3', from left to right. This applies to RNA, single-stranded DNA, and oligonucleotide sequences. This is especially important when ordering oligonucleotides. If writing the sequence of dsDNA, the convention is to write the sequence of the coding strand 5'→3', while the sequence of the template strand may be written 3'→5' such that it demonstrates its antiparallel relationship with the coding strand.
22. Do not store primers below a concentration of 20 μM. Primers diluted below this concentration are often unreliable in later experiments.
23. Do not store primers or templates at excessively high concentrations, which is an excellent way to promote amplicon anarchy (massive carryover contamination) in the lab.
24. Always keep tubes containing RNA samples on ice unless the protocol specifically dictates otherwise. When reading through the protocol, let the samples rest on ice until the next course of action has been discerned. Also, be sure to prechill centrifuge tubes when carrying out a sequential extraction regimen so that aliquots are always delivered to a cold tube. Remember that cool temperatures partially inhibit nuclease activity.
25. When a protocol dictates placing a heat-denatured sample on ice immediately after heating, it is a good idea to have an ice bucket close to the heating block or thermal cycler. Plunging the tube into ice, and then traveling back to one's work area will maintain the sample in a denatured state for as long as possible.
26. Take logistical measures to keep stock solutions of RNA concentrated. This will facilitate UV spectrophotometric measurements, and subsequent dilution, as needed. Further, it is mechanistically easier to recover small RNA yields in minimal volumes from 1.7 ml microfuge tubes than from larger tubes.
27. Do not allow nucleic acid pellets to dry out completely unless explicitly specified in the procedure. RNA and DNA pellets that are *slightly* damp with ethanol will dissolve more readily than those that are completely dry.
28. Spend the extra money and purchase molecular biology-grade reagents. The use of these highly purified, nuclease-free materials will favorably increase the yield, stability, and utility of nucleic acid preps for all downstream applications.

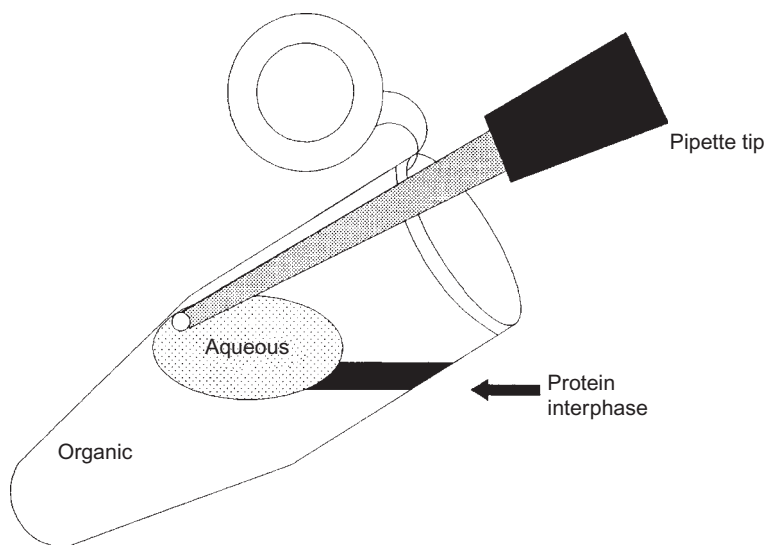


Figure 1 Technique for the quantitative recovery of aqueous material during an organic extraction. Be sure to wear appropriate eye and skin protection.

29. In addition to RNase, keep in mind that heavy metals can cause RNA degradation when they are present for extended periods. If this is a suspected problem, filtering buffers through Chelex (Bio-Rad) can help. In addition, the inclusion of 8-hydroxyquinoline, which chelates heavy metals and is itself a partial RNase inhibitor, may optimize RNA purification efficiency.
30. In the course of extraction with organic solvents to deproteinate nucleic acid samples, quantitative recovery of the aqueous phase without disturbing the protein interface is often difficult. To maximize the recovery of aqueous material after most of it has been removed, it helps to tilt the tube to a 45-degree angle, causing a “bubble” of the remaining aqueous material to form up against the side of the tube (Fig. 1). At this point, one may slide the micropipette tip along the upper inside wall of the tube, penetrate the aqueous bubble, and remove most of the remaining aqueous volume without disturbing the protein interphase or organic material. **CAUTION: Use great care when handling tubes containing organic solvents, wear appropriate eye protection and skin protection, and use common sense.**
31. When purifying nucleic acids with phenol, always perform a final extraction with chloroform alone or a mixture of chloroform:isoamyl alcohol (24:1). Because of the great solubility of phenol in chloroform, even trace amounts of carryover phenol will be removed in the final chloroform extraction. This is necessary because quinones, the oxidation product of phenol, can compromise the integrity of the nucleic acid sample.
32. Do not use polystyrene plasticware or thin-wall PCR tubes with organic solvents. The plastic will begin to dissolve in phenol or chloroform, resulting in catastrophic failure.
33. Be sure to check for β -mercaptoethanol incompatibility with ultracentrifuge tubes.
34. Sodium dodecyl sulfate (SDS) is often added to buffers to control nuclease activity. In many applications it may be desirable to eliminate it because it precipitates out of solution at cold temperatures (the sodium salt of SDS is especially insoluble).

Moreover, even traces may be incompatible with further applications, especially PCR. Other types of RNase inhibitors, such as RNasin®, may be used instead.

35. When preparing buffers and reagents that include EDTA, be sure to use the disodium salt ($\text{Na}_2\text{-EDTA}$) rather than the tetrasodium salt ($\text{Na}_4\text{-EDTA}$). This is important because of the great sensitivity of many molecular biology enzymes to salt concentration. Buffers prepared with $\text{Na}_4\text{-EDTA}$ will initially be more alkaline than those made with $\text{Na}_2\text{-EDTA}$. The subsequent addition of HCl to establish pH 7–8 will greatly increase the NaCl concentration in the final buffer, and likely have a negative influence on the outcome of an experiment.
36. When collecting nucleic acid precipitates by centrifugation, be certain not to exceed the recommended *g*-force for a particular type of tube. Attention to this detail is especially important when scaling up or scaling down RNA isolation procedures, as significant *g*-forces may be required. For smaller-scale procedures, polypropylene microfuge tubes are used almost exclusively. Properly cushioned, nuclease-free Corex glass tubes are handy for larger-scale preps because of their ability to withstand the *g*-force needed to pellet precipitated RNA.
37. The ubiquitous instruction to store labile materials in “suitable aliquots” is quite vague at best, particularly because different applications require different quantities of RNA or DNA. If the precise definition of a suitable aliquot is unknown, purified RNA is best stored in precipitated form at -80° until such a determination can be made. It is unwise to store the sample dissolved in buffer until fate of the sample is determined because repeated thawing and refreezing of RNA samples dramatically reduces the life of the material. If a redissolved sample must be stored, aliquot it and then store it at -80° so that an aliquot is thawed only once.
38. Do not discount the importance of key UV absorbance ratios as indicators of the purity of nucleic acid samples. Be sure to examine A_{260}/A_{280} , A_{260}/A_{230} , and A_{260}/A_{240} , as described in Chapter 6. Deviant ratios may be indicative of protein contamination, phenol contamination, excess salt, or incomplete solubilization of the RNA pellet.
39. Contrary to popular belief, it is not detrimental to hybridize more than one filter in the same hybridization chamber. As long as there is sufficient buffer to keep the filters from sticking together and the filters are free to move in the chamber with the motion of the incubator (orbital shaker incubators work well), hybridization of two or more filters should proceed unencumbered.
40. If a particular hybridization buffer continually produces elevated or splotchy background on the X-ray film, think about performing a sham hybridization with the hybridization buffer first (with or without the probe). In other words, incubate the hybridization buffer with a sheet of filter paper with nothing on it. The idea is that anything in the hybridization buffer that will stick nonspecifically to the filter will stick to the blank filter first. Then, transfer the hybridization buffer to the bag or container with the experimental filter. Sometimes, the improvement is astonishing.
41. Define the linear range of X-ray films when preparing to record experimental data by autoradiography or chemiluminescence. For example, preparing dilutions of at least one sample will facilitate accurate interpretation of data. Remember that changing the exposure time for X-ray film, just like altering image analysis exposure time and aperture settings digitally or with Polaroid film, will minimize problems associated with the linear response of film in general and highlight the differences among the samples.
42. To permanently record the location of size standards on X-ray film, label a very small amount of probe that will be able to hybridize to the size markers, making

detection of these nucleic acid species possible as well. This approach will eliminate guesswork regarding the precise location of molecular weight standards and comparing them to experimental samples.

43. Three-hole punch dry X-ray films and place in a three-ring notebook for organizational and storage purposes. Place a sheet of paper between films to prevent them from becoming scratched.
44. All chemicals should be considered potentially hazardous. Individuals should be trained in good laboratory technique and safety practices by their respective departments. In a molecular biology setting, there are many inherent dangers associated with what is considered “the daily routine”. Always follow the safe handling, containment, and personal safety recommendations of the manufacturers of chemicals and equipment.
45. Be sure to keep and maintain easy access to the Material Safety Data Sheet (MSDS) that accompanies hazardous chemicals and reagents. It is important for all members of a lab team to become familiar with the potential health hazards associated with each of these compounds. Each MSDS also contains appropriate first aid measures in the event of accidental exposure.
46. Do not store microfuge tubes in liquid nitrogen, ever.
47. Using a mixture of oligo(dT) and random primers goes a long way toward improving first-strand cDNA synthesis, both in terms of size and yield.
48. Failure to inactivate RT after first-strand synthesis (one-tube PCR notwithstanding) is a big no-no.
49. Never mouth pipette and always wear safety glasses.
50. Wear appropriate attire in the lab at all times: no jeans with holes, no long hair, no face piercings, no sandals, no open-toe shoes. And wear a lab coat.

Appendix A: Maintaining Complete and Accurate Records

The absolute necessity for maintaining accurate records of experiments performed and the resulting data is indisputable. Perhaps no other facet of scientific inquiry has the potential to have the greatest impact on research than the *art* of excellent record keeping. Many a graduate student has been amazed by the intralaboratory fallout that results from the failure to produce well-written, easy-to-interpret records of one's research-related activities. Not only are accurate, meaningful records an important component of continuous productivity, but also such records are often the only verification that the work has been completed and the order in which each tantalizing bit of data was generated.

Students who begin their research experience as undergraduates have a definite advantage over their graduate-student counterparts. Laboratory investigation at even the most rudimentary level, which may involve little more than “grunt” work, is often the first exposure that students have to the imperative for keeping a notebook. Although requests from more senior members of the lab may seem inane at times, invariably the time will come when that inane information must be recalled. It is the opinion of this Author that it is better to write down too much than to leave anything to the imagination.

The head of the laboratory is the party with the principle responsibility for breaking bad habits that may be imported into the lab and preventing new members of the lab from acquiring any bad habits with respect to record keeping and poor lab technique. What follows is a list of guidelines that one would reasonably expect to be adhered to in a research setting. This is not intended to be an all-inclusive list and should be modified to suit the local clientele and personalities involved.

1. Record each facet of each experiment as it is being performed. Do not trust details to memory and do not wait until the end of the day to write everything down.
2. Do not write procedural information or data on scraps of paper. All too often this type of stationery is misplaced or accidentally discarded. It is also possible that data might be transferred to record books in an incorrect order.
3. If possible, laboratory records should be maintained on a computer, and the data backed up daily.
4. Clarity is of the utmost importance. It is useless putting procedures into complex sentences garnished with elaborate punctuation if the reader is unable to understand and duplicate the procedure exactly from the notebook alone.

5. Don't worry about experiments that did not work out as expected or did not work out at all. It is imperative that everything is recorded: often, useful information can be gleaned from scrutiny of a string of failures. Just like everything else in life, we all learn from our mistakes in the laboratory.
6. A professional lab notebook is a bound journal and has a number on each page.
7. Put the date on each page of the notebook as procedures and data are entered.
8. At the beginning of each new experiment, include a sentence or two that provides a rationale for the experiment and be sure to reference or otherwise call attention to previous experiments or samples that have relevance to the business at hand.
9. Be sure to record lot numbers for each new reagent, kit, or other chemistry that is used to support the ongoing research. Many suppliers now provide all relevant data on adhesive strips on the spec sheets that accompany their products; these strips can be easily applied to the current page of one's notebook for future reference. Expiration dates should also be recorded.
10. Laboratory records should be reviewed periodically by someone associated with the project to verify the completeness of these records and the ease of comprehension. Further, suggestions as to how record-keeping might be improved can be proffered, as needed.
11. If data are manually entered into laboratory notebooks, entries should be made in black ink only. This is an industry standard. Of late, pens with fluorescent- and pastel-colored ink have become stylish; these are difficult to read, very unprofessional for everything, and should be avoided.
12. If one's handwriting is poor, then one should print instead.
13. Never "white-out" or scratch out entries in laboratory notebooks. If changes are to be made, a single line should be made through the incorrect entry, initialed by the author, and a brief expository note describing the reason for the change should be included.
14. Never tear or otherwise remove pages from a lab notebook.
15. There is no substitute for a well-motivated investigator.

Appendix B:

Converting Mass to Moles

Each investigator has a preferred method for preparing and labeling stock solutions. Some people prefer to describe solutions in $\mu\text{g}/\mu\text{l}$ (mass per volume) while others prefer to describe everything in terms of nmols (a mole quantity) or μM (molarity; e.g., $\mu\text{moles per liter}$). It is important that persons performing molecular biology methods have the ability to quickly convert mass to moles, moles to mass, $\mu\text{g}/\mu\text{l}$ to molarity, and molarity to $\mu\text{g}/\mu\text{l}$. The conversions can often be performed in one's head and without a calculator by adherence to a few simple mnemonics and remember, practice makes perfect.

With respect to PCR, the most important thing to remember is that most protocols require equimolar amounts (same number of molecules) of the upstream- and downstream primers. In other words, there is a 5' primer molecule for every 3' primer molecule. Thus, adding a $1\mu\text{l}$ of $100\mu\text{g}/\mu\text{l}$ stock solution of each primer to a master mix may not do the trick, particularly if the length and base composition of the primers differ. Therefore, primer stock solutions are best labeled in terms of their molarity. In this laboratory, primers are prepared as $200\mu\text{M}$ stock solutions, an aliquot of which is diluted down to $10\mu\text{M}$ (1:20 dilution) for routine preparation of PCR master mixes. Consequently, the primers are used at a working concentration of $0.1\mu\text{M}$, meaning that a simple 1:100 dilution of the $10\mu\text{M}$ stock is all that is required (e.g. $200\mu\text{l}$ master mix requires $2\mu\text{l}$ of the primers). Good advice: keep primer and other stock solutions at a concentration that is easily divisible by some multiple of 10.

Regarding the concentration conversions, the first and perhaps easiest thing to remember is that $\mu\text{M} = \text{pmol}/\mu\text{l}$. Second, in order to move from molar concentration to mass per volume (e.g., $\mu\text{g}/\mu\text{l}$) the molecular weight of the material must be known. Third, when performing a conversion of some type, keeping track of the units will usually prevent making a mistake. In the case of primers, which would be the most common application, the molecular weight can be determined in advance through the use of any of the free on-line tools used to predict T_m during primer design. When primers are delivered, the specification sheet that accompanies the product also provides molecular weight information.

For example, consider the oligonucleotide sequence GCTGTCGATCGATTC GATCGCTAGA, a 25-mer with a molecular weight of 7673.0 g/mol.

Scenario 1

A new tube of the primer described above arrives in the lab from the manufacturer. The spec sheet accompanying the primer indicates that the tube contains 247 μg of the primer. You add 247 μl H_2O to produce a 1 $\mu\text{g}/\mu\text{l}$ concentration. You then realize that you should express the concentration in terms of molarity. To make the conversion, perform the following calculation:

$$\frac{1 \mu\text{g}}{\mu\text{l}} \quad \left| \quad \frac{10^6 \mu\text{l}}{\text{L}} \quad \right| \quad \frac{\text{g}}{10^6 \mu\text{g}} \quad \left| \quad \frac{\text{mole}}{7673 \text{ g}} \quad \right| \quad \frac{10^6 \mu\text{mol}}{\text{mol}}$$

By multiplying all numerators and all denominators and, of course, keeping track of the units, one finds that a 1 $\mu\text{g}/\mu\text{l}$ stock solution of this primer is also has a concentration of 130 μM .

Scenario 2

You have 247 μl of a 130 μM primer stock solution and need to dilute all of it down to 100 μM in order to conform to the standard operating procedure in the lab. To make the conversion, perform the following calculation:

$$130 \mu\text{M} \text{ (initial concentration)} \div 100 \mu\text{M} \text{ (final concentration)} \\ = 1.3 \text{ (dilution factor)}$$

$$247 \mu\text{l} \text{ (initial volume)} \times 1.3 \text{ (dilution factor)} \\ = 321 \mu\text{l} \text{ (final volume to prepare)}$$

$$321 \mu\text{l} \text{ (final volume)} - 247 \mu\text{l} \text{ (initial volume)} \\ = 74 \mu\text{l} \text{ (amount of water to add)}$$

Thus, adding 74 μl H_2O to the 247 μl of 130 μM stock will produce 321 μl of a 100 μM solution.

Scenario 3

The new tube of the primer described above arrives in the lab from the manufacturer. The accompanying spec sheet indicates the properties of the primer to be 4.17 nmol per OD_{260} (optical density at 260 nm) and that the tube contains 3.7 ODs of material. You want to make a 50 μM stock solution and need to figure out how much water to add to the tube. To complete your task, perform the following calculation:

$$\frac{3.7 \text{ OD}}{\text{OD}} \quad \left| \quad \frac{4.17 \text{ nmol}}{10^3 \text{ nmol}} \quad \right| \quad \frac{\mu\text{mol}}{50 \mu\text{mol}} \quad \left| \quad \frac{\text{L}}{50 \mu\text{mol}} \quad \right| \quad \frac{10^6 \mu\text{l}}{\text{L}}$$

By multiplying all numerators and all denominators and, of course, keeping track of the units, one finds that the result is 309 μl . Thus, the addition of 309 μl to the tube described above will produce a 50 μM stock solution. Note that the contents of the fourth column pertain to the final desired 50 μM concentration.

Scenario 4

The principal investigator of the lab hands you a tube containing the oligonucleotide described above at a concentration of 150 μM and tells you to remove 20 μg from the tube for an experiment. To complete your task, perform the following calculation:

L	$10^6 \mu\text{mol}$	mol	g	$10^6 \mu\text{l}$	20 μg
150 μmol	mol	7673 g	$10^6 \mu\text{g}$	L	

By multiplying all numerators and all denominators and, of course, keeping track of the units, one finds that the result is 17.4 μl . Therefore, the 17.4 μl aliquot from the 150 μM stock solution described above contains a mass of 20 μg .

Important

Regardless of one's mathematical prowess, *always* check the final concentration spectrophotometrically whenever a dilution is made or when lyophilized material is resuspended. Measuring A_{260} is highly desirable because the actual content of primer tubes is often differs from what is indicated on the label.

Useful Molecular Weights			
Average mw of a nucleotide = 330 g/mol			
Average mw of a nucleotide base pair = 660 g/mol			
dAdenine	313 g/mol	Adenine	329 g/mol
dCytosine	289 g/mol	Cytosine	305 g/mol
dGuanine	329 g/mol	Guanine	345 g/mol
dThymine	304 g/mol	Thymine	–
dUracil	290 g/mol	Uracil	306 g/mol
Inosine	314 g/mol	Inosine	–

The molecular weights shown in the table above are approximate and refer to the corresponding nucleoside monophosphate when it is part of a polynucleotide. Add 79 g/mol for a 5' monophosphate on DNA and 159 g/mol for a 5' triphosphate on RNA. The data presented here are for estimation purposes only.

This page intentionally left blank

Appendix C: Useful Stock Solutions for the Molecular Biologist

Notes

AE Buffer

50 mM sodium acetate
10 mM EDTA
Adjust pH to 5.3; autoclave

Ampicillin (1000×)

100 mg/ml in H₂O
Filter sterilize (0.22 μm)
Store aliquots at -20°

Denaturation Buffer

For Southern blots
0.5 N NaOH
1.5 M NaCl
CAUTION: *do not autoclave*

Denhardt's Solution (50×)

1% BSA (bovine serum albumin)
1% ficoll
1% PVP (polyvinylpyrrolidone)
Filter sterilize (0.22 μm); store aliquots at -20°

DEPC-H₂O

Diethyl pyrocarbonate-treated water
See "water"

DNA (heterologous)

10 mg/ml in water.
Boil and shear with 18 gauge needle. Store at -20°
Boil aliquots again just prior to use

EDTA-Na₂ (500 mM)

Ethylenediaminetetracetic acid
Will not dissolve below pH 7.5
Adjust to final pH of 8.0 with 10N NaOH or NaOH pellets
Autoclave; store at room temperature

Notes**Ethidium bromide**

10 mg/ml in sterile water

CAUTION: *powerful carcinogen*

Store foil wrapped at 4°

GTC (also known as solution D)

Guanidium thiocyanate

4 M guanidine thiocyanate

5 mM sodium citrate, pH 7.0

0.5% sarkosyl

Store in a foil wrapped bottle at 4°

Add 100 mM β -ME to a suitable aliquot just before use

CAUTION: *extremely chaotropic agent*

Hybridization buffer

Many alternative formulations are available,
depending upon application.

Consult Chapter 13 for details.

LB (Lauria-Bertani) Medium

Per liter: 10 g tryptone

5 g yeast extract

10 g NaCl

Adjust pH to 7.4 with 1N NaOH

Autoclave; aliquot using aseptic technique

Loading Buffer (10 \times)

50% glycerol

10 mM EDTA, pH 8.0

0.25% bromophenol blue

Many alternative formulations are possible.

Store 1 ml aliquots at -20°

MOPS buffer (10 \times)

3-[N-morpholino]propanesulfonic acid.

200 mM MOPS, pH 7.0

50 mM sodium acetate

10 mM Na₂-EDTA, pH 8.0

Filter sterilize or autoclave; store at room temperature

NaCl (saturated)

Sodium chloride

Weigh sufficient NaCl to make a 6 M solution in water.

Autoclave as needed

NaOH (10N)

Add 40 g NaOH pellets to 60 ml H₂O.

Increase volume to 100 ml.

CAUTION: *highly exothermic; do not autoclave even after cooling. Extremely caustic alkali.*

Neutralization buffer

For Southern blots

0.5 M Tris, pH 8.0

1.5 M NaCl

Autoclave or filter-sterilize as needed

NP-40 Lysis buffer

140 mM NaCl

1.5 mM MgCl₂

10 mM Tris, pH 8.5

0.5% NP-40

Optional: add RNase inhibitor just prior to use

PBS

Phosphate Buffered Saline, Ca⁺⁺/Mg⁺⁺-free

Per liter: 8 g NaCl

0.2 g KCl

1.44 g Na₂HPO₄

0.24 g KH₂PO₄

Aliquot and autoclave

Phenol, saturated

See Appendix D for explanation and protocol.

Phosphate Buffer (10 mM, Na_xH_xPO₄, pH 7.0)

Electrophoresis buffer

Per liter: 39 ml 1 M NaH₂PO₄

61 ml 1 M Na₂HPO₄

Check pH.

Adjust volume to 1 liter

Aliquot and autoclave.

SDS (10%)

Sodium dodecyl sulfate

Add 10 g SDS to 80 ml sterile water.

CAUTION: *do not breathe SDS powder.*

Increase volume to 100 ml.

Store at room temperature.

Do not autoclave.

SSC (20×)

Saline sodium citrate

3 M NaCl

0.3 M Na₃-citrate

Adjust pH to 7.0 with 1 N HCl

Autoclave; store at room temperature

SSPE (20×)

Saline sodium phosphate-EDTA

3 M NaCl

0.2 M NaH₂PO₄·H₂O

0.02 M Na₂-EDTA

Adjust pH to 7.4

Autoclave; store at room temperature

TAE (50×)

Tris-Acetate-EDTA

Per liter: 242 g Tris base

100 ml 0.5 M Na₂-EDTA, pH 8.0

57.1 ml glacial acetic acid

Handle with care

Autoclave; working concentration is 1 × TAE

TBE (20×)

Tris-Borate-EDTA

Per liter: 121 g Tris base

61.7 g sodium borate

7.44 g Na₂-EDTA

Autoclave; working concentration is 1 × TBE

TE Buffer (traditional)

10 mM Tris, pH 7.5

1 mM EDTA

Autoclave; store at room temperature

Adjust pH to 8.0 for DNA storage

TE Buffer (modified)

10 mM Tris, pH 7.5

1 mM EDTA

Autoclave; store at room temperature

Adjust pH to 8.0 for DNA storage

TE-9 Buffer

500 mM Tris, pH 9.0

20 mM EDTA

10 mM NaCl

Autoclave; store at room temperature.

Tetracycline (1000×)

12.5 mg/ml in 70% ethanol

Filter sterilize (0.22 μm)

Store aliquots at -20°

Tris (1 M)

Weigh out sufficient Tris to prepare a 1 M stock solution.

Dissolve in approximately two-thirds of the required amount of water.

Adjust to desired pH. Add water to achieve final volume.

Note: Tris base and Tris·Cl will be pH-adjusted differently.
Autoclave as needed.

Tris/SDS buffer

100 mM NaCl
1 mM EDTA
10 mM Tris, pH 8.5
Autoclave; then add 0.5% SDS
Store at room temperature

Water (DEPC-treated; RNase-free)

Add DEPC to laboratory grade (deionized/distilled)
water to a final concentration of 0.05–0.1% (v/v).
Mix thoroughly on an orbital shaker for several hours or
stir rapidly on a magnetic stirrer for 30 minutes or more.
Autoclave to destroy DEPC.

CAUTION: DEPC is a carcinogen.

*Follow all of the safety instructions recommended by the manufacturer.
Do not add DEPC to any solution which contains an amine group.
See Chapter 7 for details.*

YPD Medium*Yeast culture medium*

Per liter: 10 g yeast extract
20 g peptone
20 g dextrose
Aliquot and autoclave.

This page intentionally left blank

Appendix D:

Phenol Preparation

CAUTION: Phenol and chloroform are both volatile, caustic agents. Disposable gloves, protective eyewear, and the use of a chemical fume hood are necessary safety precautions when handling these and other organic solvents.

Phenol must be purified or redistilled prior to use in any molecular biology application. Redistilled phenol is commercially available, at a very reasonable cost, from all major suppliers of molecular biology reagents. The distillation of phenol in the laboratory is a very dangerous process that should be carried out by experienced personnel only, and even then it should be aggressively discouraged. Details of in-laboratory phenol distillation can be found elsewhere (Wallace, 1987).

In this lab, the standing procedure is to purchase molecular-grade (i.e., redistilled) phenol and store it frozen in suitable aliquots, wrapped in aluminum foil; an aliquot should not be saturated until just before it is to be used for the first time.

Phenol is very light sensitive and it oxidizes rapidly. The oxidation products, known as quinones, form free radicals that break phosphodiester bonds and crosslink nucleic acids. Quinones can usually be detected by the pinkish tint they impart to phenol reagents. To reduce the rate of phenol oxidation, 8-hydroxyquinoline may be added (Chapter 2) after the phenol has been saturated with aqueous buffer. Moreover, adding an equal volume of chloroform, in the preparation of an organic extraction buffer, stabilizes the phenol, imparts a greater density to the mixture, improves the efficiency of protein removal from the sample, and facilitates removal of lipids from the RNA preparation. Very often, isoamyl alcohol is used with chloroform or with mixtures of phenol and chloroform; isoamyl alcohol reduces the foaming of proteins that would normally accompany the mechanics of most extraction procedures. When prepared in this way, a saturated phenol reagent is known as an extraction buffer. Common organic extraction buffer formulations consist of phenol:chloroform:isoamyl alcohol in a 25:24:1 ratio.

Example: Redistilled phenol is typically supplied in 500g bottles. Upon arrival in the laboratory, the phenol is melted in a 65° water bath in a fume hood. Then, using a sterile glass pipette, a maximum of 25ml is transferred into conical 50ml polypropylene tubes (*never* polystyrene), or 100ml glass bottles (50ml maximum volume). Then, the individual aliquots are sealed, wrapped in aluminum foil, and stored at -20°. This precludes the repeated

thawing and refreezing of the unused phenol. When preparing to use the phenol, an aliquot is thawed only once and can then be saturated as described here or according to the specifications of a particular procedure. If 50 ml is a convenient volume of extracting buffer to prepare, then 12.5 ml of melted phenol is aliquoted into a 50 ml tube in advance and stored frozen.

To prepare the extraction buffer, melt the aliquot and then saturate it with the addition of an equal volume of a Tris-containing buffer (described below), thereby doubling the volume in the tube to about 25 ml. This is usually followed by the addition of an equal volume (25 ml) of a mixture of chloroform:isoamyl alcohol (24:1). This will produce 50 ml of saturated extraction buffer, which is compatible with a variety of applications. At this point, 8-hydroxyquinoline can be added to a final concentration of 0.1% (w/v); otherwise, the extraction buffer can be used as is. This basic strategy can be applied to the preparation of any volume of extraction buffer. For convenience, saturated phenol and ready-to-go extraction buffer may be purchased from Sigma Chemical Company (Cat. Nos. P-4557 and P-3803, respectively).

Phenol saturation

There are two standard types of phenol saturation and several variations under each heading. One of these methods, Tris saturation, is used when buffered phenol is to be prepared and maintained as a stock of known pH in the lab or when a particular protocol specifies that a phenol-containing extraction buffer be prepared at a specific pH before use. It is important to note here that the rate of phenol oxidation increases as pH increases. Therefore, phenol buffers greater than pH 8.0 should be prepared fresh each time they are required. Saturated phenol buffers equilibrated to more acidic pH are considered stable for 4 to 5 weeks if wrapped in foil and stored at 4°. The other method, water saturation, is used when the *starting* pH of the phenol is not of critical importance. Phenol should never be water saturated unless specifically required of a particular protocol. When water saturation is used, the final pH of the organic material is usually achieved by mixing the phenol with a suitable volume of an appropriate buffer [e.g., sodium acetate, pH 4.5, as in the method of Chomczynski and Sacchi (1987); see Chapter 2]. It is judicious to prepare only the amount of extraction buffer that will be required for a given experiment; any residual material is handled according to lab policy regarding the disposal of organic waste.

The choice of pH at which to saturate phenol is strongly dependent on whether DNA or RNA is to be isolated. A principle of paramount importance is that, in a phenolic lysate, the partitioning of nucleic acids between the aqueous and organic phases is heavily pH-dependent. Only RNA remains in the aqueous phase if the pH is acidic, and both RNA and DNA are in the aqueous phase if the pH is alkaline (Brawerman *et al.*, 1972). For example, at pH <5.5, DNA is selectively retained in the organic phase and interphase, whereas RNA can be efficiently

recovered from the aqueous phase. One key advantage of this approach is the reduced nuclease activity that is observed at acidic pH. On the other hand, if DNA is the target molecule in an extraction, adjusting the pH of the extraction buffer to pH >7.5 is the industry standard. Further, near-neutral pH equilibration very effectively safeguards against excessive depurination of the sample.

The proper pH at which to equilibrate organic solvents is also directly related to whether phenol is being used alone or in combination with chloroform and/or isoamyl alcohol. For example, extraction of aqueous lysates with phenol of acidic pH will cause the depurination of RNA (and DNA, if present) and poly(A)⁺ mRNA would be lost to the organic phase. This phenomenon can be avoided by including sodium dodecyl sulfate (SDS), chloroform, or both in the lysis and/or extraction buffers.

Procedure 1: Tris saturation

CAUTION: Be sure to wear gloves and eye protection throughout.

1. Melt a suitable aliquot of redistilled phenol completely in a 65° water bath.
2. Add an equal volume of 1 M Tris, pH 8.0 (or other pH, according to the particular application). Alternatively, phenol can be equilibrated using TE buffer (10 mM Tris, pH 8.0; 1 mM EDTA), or Tris-SDS buffer (100 mM NaCl; 1 mM EDTA; 10 mM Tris-Cl, pH 8.5; 0.5% SDS). In any event, always saturate with an equal volume of aqueous buffer.
3. Carefully mix; allow the phases to separate completely. Phase separation can be accelerated by centrifuging the tubes briefly in a desktop (clinical) centrifuge.
4. Observe the lower phenol phase and ensure that it is colorless. Phenol with any coloration, especially a pinkish tint, should be discarded because of the presence of quinones. The volume of the phenol phase should be greater than that of the upper aqueous phase, due to the preferential absorption of some of the aqueous component into the organic phase.
5. Remove the upper (aqueous) phase and discard it. To the remaining phenol phase, add an equal volume of 100 mM Tris, or other Tris buffer, adjusted to the same pH as in step 2.
6. Carefully mix; allow the phases to separate completely. Phase separation can be accelerated by centrifuging the tubes briefly in a desktop (clinical) centrifuge.
7. Remove most of the upper aqueous phase and measure its pH. If the aqueous material is pH >7.5, the phenol is now considered Tris-saturated and is ready for use. If pH <7.5, repeat steps 5 and 6.
8. Store saturated phenol wrapped in foil at 4°. If desired, mix with an equal volume of chloroform:isoamyl alcohol (24:1) and/or 0.1% (w/v) 8-hydroxyquinoline.

Procedure 2: Water saturation

CAUTION: Be sure to wear gloves and eye protection throughout.

1. Melt a suitable aliquot of redistilled phenol completely in a 65° water bath.
2. Add an equal volume of sterile dH₂O.

3. Allow the phases to separate completely before use. Phase separation can be accelerated by centrifuging the tubes briefly in a desktop (clinical) centrifuge.
4. The upper (aqueous) material can be drawn off and discarded or allowed to remain in the tube with the phenol as evidence of saturation.
5. Store saturated phenol wrapped in foil at 4°. If desired, mix with an equal volume of chloroform:isoamyl alcohol (24:1) and/or 0.1% (w/v) 8-hydroxyquinoline.

References

- Brawerman, G., J. Mendecki and S.Y. Lee. (1972). A procedure for the isolation of mammalian messenger ribonucleic acid. *Biochemistry* **11**, 637.
- Chomczynski, P. and N. Sacchi. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156.
- Wallace, D.M.. (1987). Large- and small-scale phenol extractions. *Methods Enzymol.* **152**, 33.

Appendix E: Disposal of Ethidium Bromide and SYBR Green Solutions

Ethidium bromide (EtBr), SYBR Green, and similar dyes are used to stain nucleic acids in gels are very powerful mutagens (MacGregor and Johnson, 1977). As intercalating agents and nucleic acid-binding dyes, they should be treated as toxic waste; the proper handling of these materials includes preventing environmental contamination after use. Only after proper treatment can ethidium bromide and SYBR Green-tainted wastes be disposed of with minimal apprehension.

The salient issues surrounding the treatment of ethidium bromide waste have been described at length (Lunn and Sansone, 1987, 1990; Bensaude, 1988). The fairly widespread practice of adding bleach to ethidium waste is not recommended, because such treatment, while reducing the mutagenicity of ethidium bromide, converts the dye into another mutagenic compound (Quillardet and Hofnung, 1988). Instead, one of the following protocols should be adopted as standard lab policy and applied to the handling of SYBR Green as well. The following protocols suggest the handling of working concentrations of ethidium bromide (0.5–1.0 µg/ml) and SYBR Green (1–4×); at higher concentrations, these dyes can be diluted down and processed as described in Protocol 1, Protocol 2, or Protocol 3.

Protocol 1

The Extractor (Scheicher & Schuell; Cat. No. 448031) is a one-step filtration method for the removal of ethidium bromide and SYBR Green from gel-staining solutions (Fig. E.1). The device is capable of filtering 10 liters of electrophoresis buffer containing EtBr, with greater than 99% removal (Fig. E.2). The filter, an activated carbon matrix, is disposed of according to departmental guidelines, whereas the filtrate can be discarded safely down the drain. This is an extremely cost-effective method for dealing with this common waste product in the molecular biology laboratory and is used extensively in this laboratory.



Figure E.1 Extractor filtration device for the removal of ethidium bromide and SYBR Green from electrophoresis- and staining buffers. Photograph courtesy of Whatman, Inc..

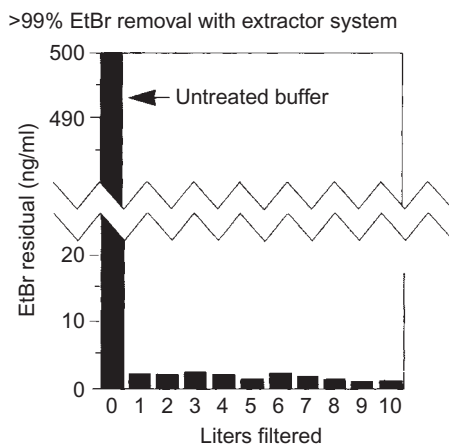


Figure E.2 Ten successive liters of buffer containing 500 ng/ml EtBr were filtered through the Extractor device. Residual EtBr in the filtrate was measured by fluorescence spectroscopy. Data courtesy of Schleicher & Schuell Bioscience.

Ancillary protocol

Quantitative Assay for Residual EtBr in Filtrate (Menozzi *et al.*, 1990).

1. Prepare a solution of 100 $\mu\text{g/ml}$ salmon sperm DNA in running buffer or staining buffer. Prepare dilutions of EtBr in the range of 0 to 500 ng/ml to prepare a standard curve.

2. Add salmon sperm DNA to 1 ml aliquots from each volume of filtrate to a final concentration of 100 µg/ml.
3. Read standards and unknowns in a fluorimeter (excitation 526 nm, emission 586 nm). Plot the standard curve, and read the EtBr concentration of the unknown from the standard curve. This assay is usually linear ($r = 0.999$) for EtBr concentrations as low as 4 ng/ml.

Protocol 2

Solutions containing working concentrations of ethidium bromide, SYBR Green, and related dyes can be decontaminated by adding about 3 g of Amberlite¹ XAD-16 (Sigma Cat. No. XAD-16) for each 100 ml of solution (Joshua, 1986; Lunn and Sansone, 1987). This resin is a nonionic, polymeric absorbent. The solution may be shaken intermittently for 12–16 h at room temperature, after which it is filtered with Whatman No.1 filter paper, or the equivalent. The filter and Amberlite should be treated as toxic waste and disposed of according to in-house laboratory policy. The filtrate can then be discarded.

Protocol 3

Add 100 mg of activated charcoal (Sigma Cat. No. C-2889) to each 100 ml of ethidium bromide or SYBR Green at the working concentration (Bensaude, 1988). The resultant mixture may be shaken intermittently for 1 h at room temperature, after which it is filtered with Whatman No.1 filter paper, or the equivalent. The filter and charcoal should be treated as toxic waste and disposed of according to in-house lab policy. The filtrate can then be discarded.

References

- Bensaude, O. (1988). Ethidium bromide and safety-readers suggest alternative solutions. Letter to editor. *Trends Genet.* **4**, 89.
- Joshua, H. (1986). Quantitative absorption of ethidium bromide solutions from aqueous solutions by macroreticular resins. *BioTechniques* **4**(3), 207.
- Lunn, G. and E.B. Sansone. (1987). Ethidium bromide: Destruction and decontamination of solutions. *Analyt. Biochem.* **162**, 453.
- Lunn, G. and E.B. Sansone. (1990). Degradation of ethidium bromide in alcohols. *BioTechniques* **8**(4), 372.
- MacGregor, J.T. and I.J. Johnson. (1977). *In vitro* metabolic activation of ethidium bromide and other phenanthridinium compounds; mutagenic activity in *Salmonella typhimurium*. *Mutation Res.* **48**, 103–108.
- Menozi, F.D., A. Michel, H. Pora and A.O.A. Miller. (1990). Absorption method for rapid decontamination of solutions of ethidium bromide and propidium iodide. *Chromatographia* **29**, 167–169.
- Quillardet, P. and M. Hofnung. (1988). Ethidium bromide and safety-readers suggest alternative solutions. *Trends Genet.* **4**, 89.

¹Amberlite resins are manufactured by Rohm and Hass, Inc. (Philadelphia, PA).

This page intentionally left blank

Appendix F: DNase I Removal of DNA from an RNA Sample

It often becomes necessary to purge RNA samples of contaminating DNA for any of several reasons, especially when preparing RNA for PCR-based applications. Although isopycnic centrifugation has been used in the past to partition grossly contaminated samples, it is time-consuming, requires expensive and highly specialized equipment, does not favor high productivity, and is not the preferred approach when the mass of both DNA and RNA in the sample totals no more than a few micrograms. Moreover, the popular acid-phenol method for separating RNA from DNA, although very efficient, is not efficient enough to ensure that only RNA remains in the aqueous environment (Chapter 2), and the closer one approaches the organic phase with the micropipette tip, the greater is the likelihood that DNA is carried over. This presents a significant compromise of the quantitative nature of RNA-based assays; probes and primers often have great difficulty distinguishing between RNA targets and DNA targets when both are present. In any of these circumstances, a brief incubation of the sample(s) with ribonuclease-free deoxyribonuclease I (RNase-free DNase I) will eliminate DNA by nuclease digestion. The resulting RNA solution can then be used directly; it may also be extracted with phenol:chloroform, and/or concentrated by standard salt and alcohol precipitation (Chapter 2), or otherwise cleaned up. Applications for DNase treatment in this manner include, but are not limited to, preparation of RNA free of contaminating DNA, degradation of DNA template molecules from *in vitro* transcription reactions, nick translation of DNA probes, and the study of DNA protein interactions by DNase I footprinting.

Commercial preparations of RNase-free DNase I, such as RQ 1 (Promega), are typically purified by any of a number of methods to remove all traces of detectable RNase and are certified by the manufacturer to harbor no intrinsic RNase activity. These preparations are very handy in the lab and well worth the purchase cost. There is no easy, cost-effective way to prepare RNase-free DNase in the lab because of the labile nature of DNase I. It is also important to note that while the following protocol calls for the use of the 10× reaction buffer that normally accompanies the purchase of DNase I, several newer kits that call for DNase treatment of a sample may not necessarily call for the use of the traditional reaction buffer. This is often permissible because the chemical environment in which the RNA is stored at a particular stage of a protocol may, in fact, be conducive to the activity of the DNase. Always verify this

important procedural step because DNase I will not function unless supported by the requisite cofactor, ionic strength, and pH requirements.

Protocol

1. To a sample of RNA suspected to be tainted with DNA, add RNase-free DNase I to a final concentration of approximately 1 U/ μ g nucleic acid.

Note 1: DNase I typically will show optimal activity in a working buffer consisting of about 40 mM Tris-Cl, pH 7.9; 1 mM CaCl₂; 5 mM MgSO₄. This or a similar formulation is generally provided by dilution of a 10 \times reaction buffer that accompanies the purchase of the enzyme.

Note 2: The activity of RNase-free DNase I is typically described in units, where 1 unit of DNase I will degrade 1 μ g of DNA in 10 min at 37 $^{\circ}$, most often performed in a 20 to 50 μ l reaction volume. Keep in mind that DNase I is strongly inhibited in buffers containing EDTA, SDS, and other denaturants or if heated above 37 $^{\circ}$.

2. Incubate the mixture for 15 to 30 min at 37 $^{\circ}$. As needed, incubations up to 1 h are also acceptable.
3. Terminate the reaction by heating the sample to 65 $^{\circ}$ for 10 min. Alternatively, the activity of DNase I can be arrested by the addition of EGTA to a final concentration of 2 mM.

Note: The addition of EDTA is not recommended for RT-PCR. Stop buffer (EGTA) is usually provided along with the enzyme.

4. Assess the extent of DNA digestion as well as the integrity of the RNA in the sample by electrophoresing an aliquot of the sample under denaturing conditions.
5. The RNA is now ready for direct use in RT-PCR.

Appendix G: RNase Incubation to Remove RNA from a DNA Sample

In some applications, such as the isolation of high-molecular-weight genomic DNA (Appendix K), the experimental design may require the removal of all contaminating RNA without compromising the integrity of the DNA. In these cases, purging of all intrinsic DNase activity from the RNase stock solution must be carried out before use. As indicated in Appendix E, isopycnic centrifugation could be employed to partition grossly contaminated samples, although it is time-consuming, requires expensive and highly specialized equipment, and is not the preferred method when the combined mass of DNA and RNA in the sample totals no more than a few micrograms. In these circumstances, a brief incubation of the mixture with DNase-free RNases will do the job. Commercially available DNase-free RNase is certified by the manufacturer to harbor no intrinsic DNase activity. Crude preparations of RNases, if not specially treated, can harbor significant levels of DNase activity as well (Gillespie and Spiegelman, 1965).

Unlike the removal of RNase activity from DNase, RNase preparations can be made DNase-free in the most basic molecular biology laboratories. The most commonly used RNases are the enzymes RNase A and RNase T1. The following procedure pertains to the preparation of the homemade version only. Commercial preparations are ready for immediate use.

Protocol: Removal of DNase activity from RNase

1. Prepare a 10 mg/ml stock solution of RNase A, RNase T1, or both in sterile H₂O or TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0).
2. Heat RNase stock solutions to near boiling (90°) for 10 min. Cool on ice and store frozen in aliquots.

Note: Although some RNase preparations maintain considerable activity after boiling, this practice is not encouraged for this particular application; at 90° DNase activity will be quickly eliminated without compromising the RNase activity of the reagent.

Protocol: Digestion of RNA

1. Add DNase-free RNase to sample to a final concentration of 10 µg/ml.
2. Incubate for 10 to 30 min at 37°.

3. Terminate RNase digestion by extraction with an equal volume of phenol:chloroform (1:1) or phenol:chloroform:isoamyl alcohol (25:24:1).
4. Concentrate DNA by salt and alcohol precipitation, or use directly for restriction endonuclease digestion, PCR, etc.
5. Store genomic DNA at 4°. Plasmid DNA can be stored for extended periods at -20°.

Reference

Gillespie, D. and S. Spiegelman. (1965). A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. *J. Mol. Biol.* **12**, 829.

Appendix H: Deionization of Formamide, Formaldehyde, and Glyoxal

Formamide, formaldehyde, and glyoxal are routinely utilized for a wide variety of molecular biology applications. The rapid oxidation of these reagents mandates deionization prior to use for optimal efficacy in hybridization and nucleic acid denaturation applications. Two common approaches for deionization are presented here.

Deionization of commonly used molecular biology reagents can be achieved by the addition of the reagent to a small mass of mixed bed resin (Sigma; Catalog No. M-8032). This mixed anionic and cationic exchange resin is strongly acidic and strongly basic, facilitating rapid deionization. Briefly, mix 50 ml of the reagent to be deionized with 5 g of mixed bed resin. Swirl the mixture every 2–3 min. As the capacity of the resin is reached, its color will change from blue to gold. Deionization should be complete in 30 min and may require the use of more than one aliquot of the resin. In the case of formamide deionization, the color change occurs rapidly, even before the deionization capacity of the resin is reached. Remove the resin by filtration or by centrifugation. Store deionized glyoxal in tightly sealed, convenient aliquots at -20° ; deionized formamide and formaldehyde should be used shortly after deionization. Once opened, the unused portion of a previously deionized aliquot should be discarded.

Alternatively, formamide, formaldehyde, and glyoxal can be deionized with AG 501-X8 mixed bed resin (Bio-Rad). AG 501-X8 resin is a 1:1 equivalent mixture of AG 1-X8 resin (OH^{-}) and AG 50W-X8 resin (H^{+}). It is recommended that 1 g of AG 501-X8 resin be added to each gram of material to be deionized. Deionization requires about 1 h, although the mixture can be left overnight. As described above, when deionizing formamide, the color change occurs rapidly, even before the deionization capacity of the resin is reached. Then the resin can be removed by filtration, and the deionized reagent stored as described above. The Bio-Rad AG 501-X8 resin is available with or without a conjugated dye that will manifest a color change as the deionization process proceeds.

This page intentionally left blank

Appendix I: Silanizing Centrifuge Tubes and Glassware

The intrinsically “sticky” nature of nucleic acids frequently mandates the silanizing of glassware (and polypropylene tubes, in some applications) for maximum sample recovery. Silanizing Corex or Pyrex glass tubes and microfuge tubes, for example, minimizes loss of RNA during larger scale isolation, and complementary DNA (cDNA) in many downstream applications, though some loss is inevitable. Silanizing is especially helpful when small quantities of sample are being manipulated.

Protocol

Sigmacote[®] (Sigma; Cat. No. SL2) is a very useful reagent for the rapid silanizing of glassware and certain types of plasticware. Sigmacote is a special silicone solution in heptane. It forms a tight, microscopically thin, water-repellent film of silicone on surfaces to which it is exposed.

1. Be sure to follow the manufacturer’s instructions for the safe handling of Sigmacote.
2. Fill articles to be coated with Sigmacote. Avoid dipping, as tubes will become very slippery if the outside becomes silanized, and it will be very difficult to write on them.
3. Drain and allow articles to air dry, which is usually complete in 5–10 min. Dry surfaces are essentially neutral.

Note: Sigmacote is reusable if kept free of moisture.

This page intentionally left blank

Appendix J: Trypsinization Protocol for Anchorage-Dependent Cells

The preparation of nucleic acids from biological material begins with cellular lysis or tissue disruption. In the case of the former, it is possible to add a lysis buffer directly to a flask or dish containing cells grown in culture. This approach, however, often necessitates a volume of lysis buffer larger than would be required had the cells been harvested from tissue culture and lysed in a separate tube. In the case of suspension cultures, cells are collected by simple centrifugation of the culture medium. For anchorage-dependent cells, the following is the protocol used routinely in this lab when preparing for RNA or DNA isolation.

Protocol

1. Decant and save growth medium from the tissue culture vessel. It is usually convenient to decant directly into a conical 15 ml or 50 ml tube.
2. Wash the cell monolayer twice with 1× phosphate-buffered saline, calcium magnesium free (CMF-PBS). Decant and discard PBS.

Note: 1× PBS (per liter) = 8 g NaCl, 0.2 g KCl, 1.15 g Na₂HPO₄, 0.2 g KH₂PO₄. If desired, concentrated stock solutions of PBS can be prepared, and diluted with sterile H₂O just prior to use.

3. Add 2 ml 1× trypsin-EDTA to each T-75 tissue culture flask, or the equivalent. Incubate for 30 to 60 sec with gentle rocking. The exact time will depend on the cell type, degree of confluency, and temperature of the trypsin. The trypsin must be removed from the tissue culture vessel *before* the cells begin to detach.

Note: 10× trypsin-EDTA = 0.5% trypsin; 0.2% EDTA, and is usually purchased, rather than formulated in the lab. Dilute 10× trypsin-EDTA stock with 1× PBS. Alternatively, 1× trypsin-EDTA can be purchased and used directly.

4. Decant and discard trypsin. Observe the rounding of the cells.

Note: In the absence of a tissue culture microscope, the rounding of cells can be assessed by holding the flask up to a window or toward the lights in the laboratory. Because the refractive index¹ changes as the cells round, the side of the flask on which the cells are attached will appear cloudy or fogged, the extent of which is dependent on the degree of cell rounding and the density of cells on the tissue culture plastic. It is extremely important that the culture not be over-trypsinized, as this may result in the premature lysis of cells and significant loss of RNA or DNA.

¹The refractive index is the angle at which light bends when it moves from air through another medium, in this case, cytoplasm.

5. Strike the flask sharply against the palm of the hand in order to dislodge the cells from the substratum. If cells fail to detach completely, warm the flask in the palm of the hand or place the flask in a 37° incubator for an additional 60s. Strike the flask again, to ensure detachment.

Note: It should not be necessary to strike the flask more than three or four times to completely dislodge the cells. If some cells appear to remain attached, longer contact with trypsin was required. A cell culture can be trypsinized again, but only after washing it with PBS. Moreover, if re-trypsinization is necessary, be sure to save the growth medium supernatant that was added back to the flask after the first attempt at trypsinization because it is likely that large numbers of cells have become dislodged from the substratum. Cells in suspension should be stored on ice until the entire cell population is trypsinized. The cells may then be pooled into a common tube.

6. Immediately after the cells have detached from the substratum, add 5 ml of medium saved from step 1 back into the flask to inactivate residual trypsin.

Note: Some cell cultures exhibit exquisite sensitivity to trypsin. In these cases, the cells would benefit from trypsin inactivation by the addition of a specific inhibitor of trypsin after the cells have detached. This is also true if the detached cells are to be resuspended in serum-free medium. Consult the literature for information about specific cell lines and medium formulations.

7. Transfer the suspension to a suitable centrifuge tube and pellet cells by centrifugation at 150–200 × g for 3–5 min.
8. Following centrifugation, be sure to remove the supernatant as completely as possible. This is necessary to ensure that the key components of the lysis buffer are not diluted or rendered useless by interaction with residual culture media. If cell growth media removal is problematic, one may wash the cell pellet with PBS and then repeat steps 7–8 as necessary.
9. To facilitate cell resuspension in lysis buffer, or any other reagent, it is very helpful to tap the tube containing the cells *after* the supernatant from step 8 has been removed, and before the addition of the lysis buffer. In this lab, tubes containing cell pellets are dragged along the top of a microfuge rack two or three times to help break up the cell pellet, immediately after which the lysis buffer is added.

Appendix K: Isolation of High-Molecular-Weight DNA by Salting-Out

The following is a modification of the procedure of Miller *et al.* (1988) for the isolation of high-molecular-weight (HMW) DNA without the use of toxic organic solvents. Although other procedures have been described for the recovery of HMW DNA, including dialysis (Longmire *et al.*, 1987) and the use of filters (Leadon and Cerutti, 1982) and some of the kits that are widely available from biotech vendors are based on these methods or the method described here.

This procedure is relatively brief and involves the salting-out of cellular proteins; it begins with peptide hydrolysis by proteinase K, followed by protein precipitation with a saturated NaCl solution. Protein-free genomic DNA is subsequently recovered by standard salt and ethanol precipitation. This approach has been used successfully for DNA isolation from nucleated blood cells (Miller *et al.*, 1988) and, in this laboratory, from a variety of trypsinized cell types of mesenchymal origin (Farrell, 1991–2009; unpublished data). In this laboratory, the resulting DNA is routinely 10 to 15% larger, on average, than genomic DNA purified with phenol:chloroform. DNA purity, as assessed by scanning UV spectrophotometry and examination of A_{260}/A_{280} (Chapter 6), is consistently comparable to that routinely achievable with phenol:chloroform protein extraction techniques.

Protocol

1. Harvest cells and pellet them in a 15 ml centrifuge tube by centrifugation at $200 \times g$ for 5 min. Use 1 to 2×10^7 cells per tube.
2. Decant supernatant. Completely resuspend cell pellet (up to 2×10^7 cells) in a total volume of 4.5 ml TE-9 lysis buffer (500 mM Tris-Cl, pH 9.0; 20 mM EDTA; 10 mM NaCl).

Optional: Wash cell pellet with PBS¹ and centrifuge again before addition of lysis buffer.

Note 1: The alkaline pH of this lysis buffer is desirable because it will induce partial hydrolysis of RNA in the sample. It is not sufficiently alkaline, however, to cause severe denaturation of double-stranded DNA (Chapter 8).

¹1 × PBS (per liter) = 8 g NaCl, 0.2 g KCl, 1.15 g Na₂HPO₄, 0.2 g KH₂PO₄. If desired, concentrated stock solutions of PBS can be prepared, and diluted with autoclaved H₂O just prior to use.

Note 2: For fewer starting numbers of cells, the volumes of all reagents should be scaled down proportionally.

3. Add 500 μ l 10% sodium dodecyl sulfate (SDS) to the lysate. Invert tube sharply to mix.
4. Add 125 μ l 20 mg/ml proteinase K (0.5 mg/ml final concentration) to the lysate and invert tube sharply to mix. Incubate at 48° for 4–20 h.

Note: The inclusion of proteinase K is necessary to at least partially hydrolyze the nucleosome proteins. If not removed, histones and other nuclear proteins will copurify with the DNA and will interfere with restriction enzyme digestion of the DNA and other downstream applications.

5. At the end of the incubation period, add exactly 1.5 ml saturated NaCl solution (approximately 6 M). Shake vigorously for 15 s. Allow the tube to sit on the bench for 1–2 min in order to observe the appearance of the lysate. When performed properly, precipitated polypeptides will cause the lysate to appear very cloudy and remain so.

Note: Additional small aliquots of saturated NaCl may be needed to cause the salting-out of polypeptides, which is essentially a dehydration technique. This will become evident if the lysate clears within 1 min after shaking the tube. If additional salt is required, add a 200 μ l aliquot, shake, and observe the appearance of the lysate. If required, repeat with additional aliquots of saturated NaCl solution until the lysate remains cloudy. Do not be overly zealous with the use of the saturated NaCl solution, however, as a point will be quickly reached where solid NaCl begins to come out of solution, resulting in a significant decrease in DNA yield.

6. Centrifuge the tube containing the lysate at 1000–1500 \times g for 10 min. Be sure not to exceed the maximum recommended g-force for the centrifuge tubes in use.
7. Decant the supernatant into a fresh 15 ml tube to dislodge any residual protein that was trapped by the foaming of the SDS. Centrifuge for an additional 10 min.
8. Decant the supernatant (which contains the DNA) into a fresh conical 50 ml tube. Avoid disturbing the protein pellet, if any. Discard the protein pellet.
9. Add 2 volumes of room temperature 95% ethanol to the tubes containing the supernatant from step 8. Swirl the tube or invert *gently* until the precipitated DNA is visible.

Note: Room temperature ethanol is recommended here because only DNA will precipitate out immediately; ice cold ethanol will drive the precipitation of some of the residual RNA as well.

10. Recover DNA using a silanized Pasteur pipette (Sigmacote is very helpful; see Appendix H). Make sure the DNA is not aspirated into the Pasteur pipette beyond the level that is silanized. Alternatively, precipitated DNA can be collected by centrifugation at 500 \times g for 1 min. In either case, transfer the DNA to a microfuge tube.

Note: In some cases, the tips of Pasteur pipettes can be heated and bent (in a Bunsen burner flame) to resemble a hook and used to fish the DNA out of solution. Because DNA sticks tenaciously to untreated glass surfaces, silanizing is essential.

11. Wash the DNA three times with 500 μ l aliquots of 70% ethanol to rinse away residual salt. A final wash with 95% ethanol will accelerate the required drying of the DNA pellet. Do not allow the DNA to dry out completely.

12. Dissolve the DNA completely in 100 to 200 μ l TE buffer (10 mM Tris-Cl, pH 7.5, 1 mM EDTA) before using the DNA for any purpose, including concentration and purity determinations. Dissolved samples of genomic DNA can be stored for extended periods at 4°.

Note 1: Warming the sample to 45 to 50° will help redissolve the DNA. Incubation at 37° may promote DNase activity, if present. In this lab, purified samples of genomic DNA are allowed to redissolve overnight at 4°.

Note 2: In addition to UV analysis, the integrity of the sample should be assessed by electrophoresing 250 to 500 ng of the sample on a minigel along with λ -Hind III size standards.

13. If necessary, the sample can be treated with DNase-free RNase, and the DNA reprecipitated, to remove any contaminating RNA.

14. Store purified genomic DNA at 4°. Freezing may compromise the integrity of the sample.

References

- Leadon, S.A. and P.A. Cerutti. (1982). A rapid and mild procedure for the isolation of DNA from mammalian cells. *Analyt. Biochem.* **120**, 282.
- Longmire, J., K. Albright, A. Lewis, L. Meincke and C. Hildebrand. (1987). A rapid and simple method for the isolation of high molecular weight cellular and chromosome-specific DNA in solution with the use of organic solvents. *Nucleic Acids Res.* **15**, 859.
- Miller, S.A., D.D. Dykes and H.F. Polesky. (1988). A simple salting out procedure for extracting DNA from nucleated cells. *Nucleic Acids Res.* **16**(3), 1215.

This page intentionally left blank

Appendix L: Electrophoresis: Principles, Parameters, and Safety

Electrophoresis is the technique by which mixtures of charged macromolecules, including proteins, nucleic acids, and carbohydrates, are rapidly resolved in an electric field. Unlike the amphoteric nature of proteins, where net charge is determined by the pH of the milieu, nucleic acids exhibit a net negative charge at any pH used for electrophoresis. In particular, the electrophoretic chromatography of nucleic acids has become a standard tool by which to characterize a sample qualitatively and quantitatively and to assay its purity. Electrophoresis is efficient and quick, offering enhanced resolution by comparison with other cumbersome procedures such as density gradient centrifugation, in which RNA can be size-fractionated for purposes other than blot analysis. That nucleic acids can be observed directly by staining the gel permits visual inspection of the progress of the separation; it also affords a reliable assessment of the integrity of the sample, and in most instances the investigator can recover specific bands in a variety of ways for subsequent characterization. Thus, electrophoresis is also a powerful technique for concentrating like-species of RNA and DNA: the parameters governing electrophoresis favor co-migration of molecules of equivalent size.

Historically, the earliest applications of this technique involved macromolecular chromatography in sucrose. Refinements in the methodology extended electrophoresis into starch gels and finally led to the implementation of polyacrylamide, in which the structure of the gel itself influences the migration characteristics of the molecules under investigation. It was quickly realized that gels with large pores were required to efficiently electrophorese high-molecular-weight nucleic acids. At first, very low percentage polyacrylamide gels (2.5%) were utilized, though the instability and unreliability of such low percentage polyacrylamide gels precluded their widespread use. Agarose, a linear polysaccharide extracted from seaweed, was later added to acrylamide gels (Peacock and Dingman, 1968; Dahlberg *et al.*, 1969) to give them enhanced physical strength. Presently, gels that consist of either agarose or polyacrylamide are favored for nucleic acid separations, depending on the size range of molecules within a sample. Currently, the two standard matrices for the electrophoretic chromatography of nucleic acids are agarose and polyacrylamide slabs, the electrophoresis of RNA usually being conducted under denaturing conditions. Less frequently, RNA may be characterized by two-dimensional electrophoresis (see DeWachter *et al.*, 1990, for review).

Theoretical Considerations¹

The following relationships describe the theoretical behavior of a charged molecule in sucrose; in actual practice other parameters delineated here factor into the exact behavior of a charged molecular species in a gel matrix. What follows is by no means intended to provide a comprehensive lesson on the subtleties of electricity, but rather is intended to review the basics necessary to understand how and why electrophoresis works.

When a molecule is placed in an electric field, the force exerted on it (F) is dependent upon the net charge of the molecule (q) and the strength of the field (E/d) into which it is placed. Therefore:

$$F = E/d \cdot (q) \quad (\text{Eq. 1})$$

in which E is the potential difference between the electrodes and d is the distance between them. In reality, a frictional force or drag opposes the migration toward an electrode. This frictional force (F) is dependent upon the size and shape (radius, r) of the molecule and the viscosity of the medium (η) through which it passes; v is the velocity at which the molecule is moving:

$$F = 6\pi r\eta v \quad (\text{Eq. 2})$$

By combining Eqs. (1) and (2), we see that

$$E/d \cdot (q) = 6\pi r\eta v \quad (\text{Eq. 3})$$

and by algebraic rearrangement one can express the velocity with which a molecule moves in an electric field as

$$v = (Eq)/d6\pi r\eta \quad (\text{Eq. 4})$$

Take-home lesson 1. What may be distilled from Equations 1–4 is that the velocity at which a molecule moves is proportional to the field strength and net charge and inversely proportional to the size of the molecule and the solution viscosity (stiffness of the gel).

Knowledge of two parameters of electricity is fundamental to understanding the mechanics of electrophoresis. In the first, described by Ohm's Law, the electrical current (I , amperes or amps) is directly proportional to the voltage (E , volts) and inversely proportional to resistance (R , ohms). Thus,

$$I = E/R \quad (\text{Eq. 5})$$

¹Adapted, in part, from Hoefer Pharmacia Biotech.

In the second, power (P , watts), a measure of the amount of heat produced, is the product of voltage (E , volts) and current (I , amps), and can be expressed as

$$P = EI \quad (\text{Eq. 6})$$

Substitution of E with the product IR from Eq. (5) permits expression of Eq. (6) as

$$P = I^2R \quad (\text{Eq. 7})$$

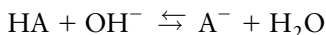
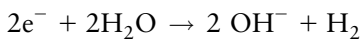
In electrophoresis, one electrical parameter, voltage, current, or power is always held constant. The consequences of an increase in resistance during the run (due to electrolyte depletion, temperature fluctuation, etc.) differ in the following ways:

- When the constant current mode is selected, in which case velocity is directly proportional to the current, heat is generated and the velocity of the molecule is maintained.
- When the constant voltage mode is selected, there is a reduction in the velocity of the charged molecules, though no additional heat is generated during the course of the run.
- When the constant power mode is selected, there is a reduction in the velocity of the molecules, but there is no associated change in heating.

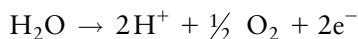
Take-home lesson 2. Temperature regulation is a primary concern throughout the electrophoretic process, particularly when running an agarose gel. Overheating will cause the gel matrix to begin to melt in an astonishingly short time, so maintaining a constant temperature is very important.

The reactions that permit the passage of current from the cathode to the anode are as follows:

Cathode reactions



Anode reactions



These reactions describe the electrolysis of water in the electrophoresis buffer, resulting in the production of hydrogen at the cathode and oxygen at the anode. For each mole of hydrogen produced, one-half mole of oxygen is

produced. By direct observation of the anode, one can see about half as many gas bubbles as are produced at the cathode. While this is certainly not the best way to ascertain whether the electrodes have been connected properly, it is certainly a simple way, as well as an indicator that current is flowing.

A typical electrophoretic separation

The electrophoresis of RNA can be executed in any of several ways although there are several features that all of these variations share. Typically, RNA is denatured in a relatively small volume (20 μ l or less), applied into the wells of a horizontal agarose gel slab, and electrophoresed for an experimentally determined period. Of course, the parameters that typically govern the electrophoresis of macromolecules apply here also. Several factors influence the electrophoretic mobility of nucleic acids (Rodbard and Chambach, 1970); once the optimal parameters have been determined in a particular laboratory setting, establishing a standard experimental format is likely to favor the reproducibility of data.

Choice of matrix

Polyacrylamide gels are typically poured between glass plates, are much thinner than agarose gels, and offer greater resolving power; agarose gels, however, allow the investigator to separate nucleic acid molecules over a much greater size range (Table L.1). This is true because the pores of an agarose gel are large, permitting efficient separation of macromolecules such as nucleic acids, large proteins, and even protein complexes. Polyacrylamide, in contrast, makes a small pore gel and can be used to separate smaller- to average size proteins and oligonucleotides. In general, the stiffer the gel, that is, the more agarose or polyacrylamide makes up the matrix, the greater is its resolving power. Agarose gels are used to electrophorese nucleic acid molecules as small as 150 bases to more than 50,000 bases (50kb), depending on the concentration of the agarose and the precise nature of the applied electric field (constant or pulse-field). Polyacrylamide gels, in contrast, are routinely used to separate nucleic acid molecules as small as 5 to 10 bases to as many as 500 to 600 bases. The resolving power of such gels is evident when one considers that polyacrylamide gels used for sequencing resolve nucleic acid molecules that differ in size by one base. **CAUTION:** *Monomeric acrylamide, that is, acrylamide in the unpolymerized form, is a very potent neurotoxin, and the assistance of someone skilled in the preparation of gels and handling of acrylamide should be sought.* Conveniently, premade gels may now be purchased from a number of suppliers and are available for most standard gel box designs and formats. Alternatively, working with agarose, which is essentially non-toxic, facilitates ease in handling, gel preparation, and disposal.

Table L.1 Useful Range of Agarose and Polyacrylamide Gels in the Electrophoresis of Nucleic Acids^a

Matrix type	Percentage (w/v)	Useful range (bases)
Agarose ^b	0.6	1,000–20,000 ^c
	0.8	600–10,000
	1.0	500–9,000
	1.2	300–6,000
	1.5	200–3,000
Acrylamide ^d	3.5	100–1,500
	5.0	80–500
	8.0	60–400
	12.0	40–200
	15.0	25–150
	20.0	6–100

^aTable derived, in part, from data presented by Sambrook and Russell (2001); Ogden and Adams, (1987); and Farrell, (1993).

^bAgarose gels consisting of as much as 3% agarose may be used to observe short products (100 to 500 bp) of the polymerase chain reaction.

^cLarger DNA molecules can be resolved efficiently by pulse-field electrophoresis.

^dGel composition = acrylamide : N,N'-methylene-bis-acrylamide, 19:1.

Polyacrylamide gels may be cast in a variety of different shapes; agarose gels, however, are traditionally poured and electrophoresed as a horizontal slab. This configuration is desirable because the gel is fully supported by the casting tray beneath it. Some vertical apparatuses accommodate the weaker nature of agarose gels by the inclusion of a frosted glass plate to which the gels may cling. Polyacrylamide, by virtue of its much greater physical strength, is almost always electrophoresed in a vertical configuration, cast either as a slab or as a cylinder and usually mounted between two buffer chambers containing separate electrodes, so that the only electrical connection between the two chambers is through the gel. In contrast, horizontally positioned agarose gels are completely submerged in electrolyte buffer, albeit with the minimum volume necessary to completely cover the gel. This facilitates reasonable dissipation of heat when the gel is run at room temperature and at sufficiently low voltages.

When the experimental design calls for the blotting or transfer of RNA out of the gel after electrophoresis, as in the traditional Northern analysis, the matrix of choice is agarose, from which RNA (and DNA for Southern analysis) can be blotted efficiently and completely. Moreover, many of the classical problems associated with preparative agarose gel electrophoresis, namely coelution of contaminants from the agarose and difficulty in elution, have been circumvented with the widespread use of extremely high-purity and low-melting-temperature agarose formulations.

Purity of agarose

Agarose is a highly purified polysaccharide derived from agar. As with all commercially available reagents for molecular biology, there are several grades, many of which contain at least trace amounts of anions such as pyruvate and sulfate (which can contribute to electroendosmosis). One should only use very high-quality molecular biology-grade agarose, certified to be low in other contaminants, including polysaccharides, and salts, and completely free of nuclease activity. These contaminants may interfere with the subsequent manipulations of nucleic acids (especially DNA) upon recovery from the gel. Highly purified grades of agarose, such as SeaKem and NuSieve (Cambrex Bio Science), are free of nuclease activity as well as most of the aforementioned contaminants. Special low-melting- or low-gelling-temperature agaroses are also available for specialized applications, and the reader is referred to the publication “A Handbook for Gel Electrophoresis” (2003). This guide, published by Cambrex (800-521-0390), is an excellent laboratory resource.

Agarose concentration

The concentration of agarose (the “stiffness” of the gel) profoundly influences the rate of migration; the mobility of a molecule of any given size differs as the concentration of agarose changes. This is because the pore size of an agarose gel is determined by the concentration of the agarose in the gel. The migration rate is linear over most of the length of the gel, although very large or very small molecules may end up outside its linear range. If this occurs, significant error may result from size determinations based on comparisons to size standards that fall *within* the linear portion of the gel. In addition, there is a certain amount of smearing of low-molecular-weight species frequently observed at the dye front (the so-called “leading edge” of the gel). This is especially noticeable in gels made up of less than 1% agarose. The usual working concentration of agarose is between 0.6% and 1.5%, although higher or lower concentrations may be adapted to specific circumstances, e.g., the use of 3% agarose gels for the analysis of PCR products.

Polyacrylamide concentration

Polyacrylamide gels offer much greater tensile strength than agarose gels. In forming the gel, acrylamide monomers² polymerize into long chains that are covalently linked by a crosslinker. It is the crosslinker that actually holds the

²**CAUTION:** Monomeric acrylamide is a potent neurotoxin and should be handled with extreme care, according to all safety precautions recommended by the manufacturer.

structure together. The most common crosslinker is N,N'-methylene-bis-acrylamide, or **bis** for short. Other crosslinkers whose special properties aid in solubilizing polyacrylamide are also used occasionally.

The pore size in a polyacrylamide gel may be predetermined by (1) adjusting the total percentage of acrylamide; or (2) varying the amount of crosslinker added to the acrylamide to induce polymerization. When there is a wide range in the molecular weights of the material under investigation, the researcher may prepare a pore gradient gel in which the pore size is larger at the top of the gel than at the bottom. Thus, the gel becomes more restrictive as the electrophoretic run progresses.

Molecular size range of sample

Nucleic acid molecules migrate at a rate inversely proportional to the \log_{10} of the number of bases (base-pairs for double-stranded DNA and DNA:RNA hybrids) that make up the molecule. This is true because the larger the molecule, the greater the frictional coefficient, or drag, that these molecules experience (Helling *et al.*, 1974) as they are pulled through the pores that constitute the geometry of the matrix. A size calibration curve for electrophoretically separated RNA can be easily constructed by plotting the \log_{10} nucleotides of at least two molecular weight standards against the distance (in centimeters) that these standards migrated from the origin (the well into which they were loaded). This is routinely performed through the application of image analysis software.

Nucleic acid conformation

The mobility of DNA is influenced greatly by the topology of the molecule (see Sambrook and Russell, 2001 for details). The electrophoretic behavior of RNA is also influenced/modified by the presence of secondary structures, also known as hairpins, which non-denatured RNA molecules frequently assume. To force identical species of RNA to comigrate, the electrophoresis of RNA is routinely carried out under denaturing conditions (Chapter 9).

Applied voltage

The efficient dissipation of heat is directly related to the resolving power of a gel. When working with agarose gels, one must be certain that the applied voltage does not result in the overheating of the gel. Gels run at relatively low voltage can be cooled as a direct function of gel box configuration. For example, vertical gels tend to dissipate heat more rapidly than horizontal gels. Further, gel boxes with an aluminum backing conduct heat many times more rapidly than glass of

a comparable thickness. In some cases, it may be necessary to continually pump cool water around the electrophoresis chamber to maintain adequately low temperatures and generate flat, high-resolution bands, and equipment so designed generally is outfitted with the required plumbing. Gels run at excessive voltages will manifest a noticeable loss of resolution and the gel will actually melt if the system becomes too hot. Agarose gels should never be electrophoresed above 5 V per centimeter of distance between the electrodes; of course gels may be run at much lower voltages. In general, the slower the gel runs, the better the resolution (2–10 kb range). If agarose gel electrophoresis is maintained within these parameters, it should not be necessary to cool the system.

Ethidium bromide

CAUTION: *Ethidium bromide is a powerful mutagen. Handle and dispose of it using great care and following the directions of the manufacturer. See Appendix E for disposal details.*

Unlike DNA, which naturally exists in double-stranded form and the migration of which can be profoundly affected by the presence of ethidium bromide during electrophoresis, RNA structure is not intrinsically influenced by the addition of ethidium bromide to an agarose gel and running buffer. When used, ethidium bromide (10 mg/ml stock) is most often diluted to a final concentration of 0.5 $\mu\text{g}/\text{ml}$. The dye does, however, retard the rate of sample migration by about 15%. It is usually just as easy and far less messy to stain (and destain) the gel at the conclusion of electrophoresis. Of even greater concern is the fact that ethidium bromide can interfere with the transfer efficiency of RNA and DNA during Northern blotting and Southern blotting, respectively. If ethidium bromide is added directly to the gel during preparation, be sure to add it to the electrophoresis buffer, too, because as the RNA runs from – to +, the ethidium bromide runs from + to –, thereby depleting the dye from the gel. Adding ethidium bromide to the electrophoresis buffer is not necessary, however, if the electrophoresis is short, for example, in checking for size, mass, or integrity. Problems occur only when the nucleic acid and the ethidium bromide pass each other on the way to opposite electrodes. For a more thorough discussion of the merits of ethidium bromide as well as other staining alternatives, see Chapter 9.

SYBR[®] Green

CAUTION: *SYBR Green, although not believed to be as mutagenic as ethidium bromide, should be handled with the same caution as any nucleic acid binding dye. Handle and dispose of it using great care and following the directions of the manufacturer. See Appendix E for disposal details.*

Three relatively new dyes, SYBR Green I, SYBR Green II, and SYBR Gold (Invitrogen) have become quite popular for use in nucleic acid electrophoresis. The SYBR dyes offer significant advantages over ethidium bromide, such as greater sensitivity, low background, and reduced mutagenicity; these differences are discussed in greater detail in Chapter 9.

It is not advisable to add any of the SYBR dyes to molten agarose when preparing the gel because of the negative effect that the dye has on the migration pattern of both DNA and RNA. Gels run in the presence of SYBR Green show bands that are fuzzy, highly irregular, and often indistinguishable from other bands in the vicinity. This phenomenon seems to occur at all concentrations of SYBR Green and at all voltages. In this laboratory, samples are electrophoresed first, followed by staining with either SYBR Green or SYBR Gold. When staining is performed after electrophoresis, the results are outstanding.

Base composition and temperature

Unlike polyacrylamide gels, the migration characteristics of nucleic acids in agarose are not influenced by their respective base compositions. Further, the electrophoretic behavior of nucleic acids in agarose gels does not change appreciably between 4° and room temperature, although the overheating of gels is a matter of great concern. Nucleic acid samples are routinely electrophoresed in agarose gels at room temperature.

Field direction

Conventional electrophoresis is accomplished by separating nucleic acid molecules in an electric field of constant direction. The classical size limitations associated with electrophoretic methods in the past have been overcome through the use of pulse-field gel electrophoresis (Schwartz *et al.*, 1982). This technique is performed by forcing DNA molecules to periodically reorient from one electric field direction to another. This approach permits efficient, reproducible separation of DNA molecules ranging in size from 0.1 to 10 megabases, and even entire chromosomes. The technology of alternating field direction is not commonly applicable to the electrophoresis of RNA, which naturally exists in the form of significantly smaller molecules.

Types of gel boxes

Agarose gels historically have been cast and electrophoresed as horizontal slabs principally because of the low tensile strength that these gels exhibit. The horizontal configuration results in the full support of the gel from beneath while

facilitating efficient separation of a wide range of nucleic acid samples, through relatively low percentage gels. Although gel box chambers of diverse designs and quality are universally available, it is definitely worth investing in a solid, durable gel box that will withstand normal laboratory wear and tear for several years. Because agarose does not exhibit the toxic qualities associated with monomeric acrylamide, it is very easy to cast an agarose gel on the bench top (or in the fume hood if toxic denaturants are added) and run the gel.

In practice, the open ends of the gel casting tray must be sealed to retain molten agarose until the matrix has solidified. Some newer designs incorporate built in rubber gaskets that seal the casting tray when inserted into the chamber 90 degrees to the intended direction of electrophoresis (Fig. L.1). This design precludes the taping requirement and frequent leaks associated with older style trays.

Better value electrophoresis gel boxes (tanks) come with a variety of combs that can be used to generate different numbers of wells of different sizes. The teeth of typical combs generally are 1-, 1.5-, 2-, or 3-mm thick and are capable of generating 10 to 20 wells per gel. It may also be desirable to purchase gel boxes with buffer outlets or ports which can accommodate buffer circulation. For example, it is critical to prevent the formation of a pH gradient when performing electrophoresis using phosphate or another buffer that does not at all resist changes in pH, such as when electrophoresing glyoxylated RNA (Chapter 9).

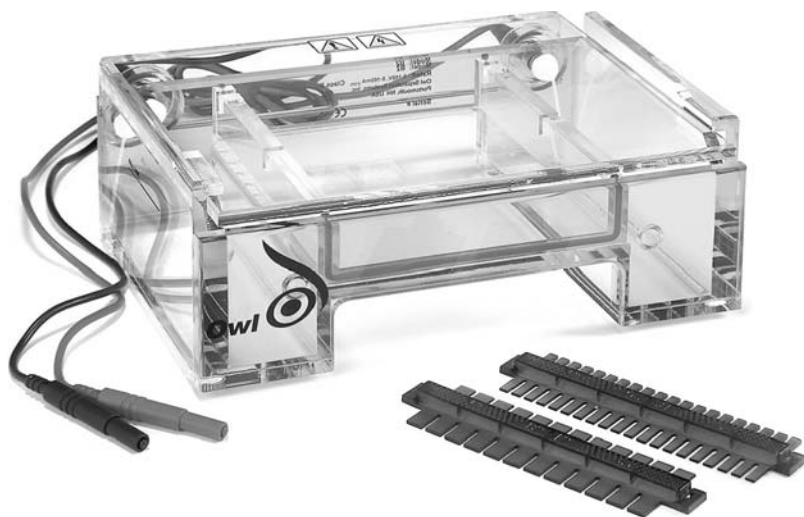


Figure L.1 Medium-size gel electrophoresis chamber featuring gasketed tray. Note the positioning of the tray and combs. After the agarose has solidified, the casting tray is rotated 90 degrees to align the wells formed by the comb in the direction of the electric field. Courtesy of Owl Separation Systems.

Safety considerations in electrophoresis

Safety must also factor into the selection of an electrophoresis apparatus. Never attempt to manipulate the gel, gel box, or power supply while conducting electrophoresis or while the power supply is plugged into an electrical outlet. Nor should any attempt ever be made to modify the design of either a gel box or power supply; any alterations could present a serious risk to both the operator and unsuspecting colleagues in the laboratory. Well-designed systems make it impossible to open or remove the lid while maintaining the flow of electrical current. In some designs the lid slides off, breaking electrical contact before the interior of the tank is accessible. This permits the safe inspection of the progress of the electrophoresis: The lid must be replaced to restore electrical connections and resume electrophoresis. Moreover, it is wise to disconnect the leads from the power supply as well as the gel box when accessing the tank for an inspection of the system. Many of the late-model power supplies are equipped to accommodate two sets of leads from two different gel boxes. Be sure to leave the power supply off until the leads have been connected to the power supply and the gel box; be sure to turn the power supply off before connecting a second set of leads (another gel box) to a power supply that is already operational. After all connections have been made properly, the power supply may be turned on again. Be sure to turn the power supply off before attempting to disconnect any of the leads. Last, *never* take hold of both leads simultaneously; there is a possibility that he who does so may complete the circuit, resulting in electrocution.

Maintenance of electrophoresis equipment

Electrophoresis is literally a daily ritual in most molecular biology laboratories. While good quality electrophoresis gel boxes and power supplies can represent an investment of thousands of dollars, it is an unfortunate, if not potentially dangerous, practice to ignore the upkeep of the instrumentation. Basic preventative maintenance of electrophoresis apparatuses is cost effective and ethically compelling.

The proper maintenance and use of electrophoresis gel boxes and power supplies is an important, though often overlooked part of lab safety. It is clearly the responsibility of the operator to inspect wires, connections, and electrodes on a regular schedule, to prevent potentially life-threatening injuries.

Perhaps the most common problem is the exposure of bare copper wire at the junction of the power cords, electrode leads, and plug ends. This situation poses a severe shock hazard. Replacement cords are usually available at nominal cost from the manufacturer or distributor and it is well worth ordering a few extras when the equipment is initially purchased. In addition, all power cords, insulation, electrodes, connection nuts, and gaskets should be scheduled

for intra-laboratory inspection on a weekly or biweekly basis. For a detailed regimen for electrophoresis apparatus maintenance, see Landers (1990).

References

- Dahlberg, A.E., C.W. Dingman and A.C. Peacock. (1969). Electrophoretic characterization of bacterial polysomes in agarose-acrylamide composite gels. *J. Mol. Biol.* **41**, 139.
- DeWachter, R., J. Maniloff and W. Fiers. (1990). Two-dimensional gel electrophoresis of nucleic acids. In *Gel Electrophoresis of Nucleic Acids – A Practical Approach* (D. Rickwood and B.D. Hames, Eds). IRL Press, Washington, DC. pp. 151–200.
- Farrell, Jr., R.E. (1993). *RNA Methodologies. A Laboratory Guide for Isolation and Characterization*. Academic Press, San Diego, CA.
- Helling, R.B., H.M. Goodman and H.W. Boyer. (1974). Analysis of R. *EcoRI* fragments of DNA from lambdaoid bacteriophages and other viruses by agarose-gel electrophoresis. *J. Virol.* **14**, 1235.
- Landers, T. (1990). Electrophoresis apparatus maintenance. *Focus* **12**(2), 54.
- Ogden, R.C. and D.A. Adams. (1987). Electrophoresis in agarose and acrylamide gels. *Methods Enzymol.* **152**, 61.
- Peacock, A.C. and C.W. Dingman. (1968). Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose-acrylamide composite gels. *Biochemistry* **7**, 688.
- Rodbard, D. and A. Chambach. (1970). Unified theory for gel electrophoresis and gel filtration. *Proc. Natl. Acad. Sci.* **65**, 970–977.
- Sambrook, J. and D.W. Russell (Eds). (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Schwartz, D.C., W. Saffran, J. Welsh, R. Haas, M. Goldenberg and C.R. Cantor. (1982). New techniques for purifying large DNAs and studying their properties and packaging. *Cold Spring Harbor Symp. Quant. Biol.* **47**, 189.

Appendix M: Polyacrylamide Gel Electrophoresis

CAUTION: Acrylamide in the unpolymerized (monomeric) form is a very potent neurotoxin that is absorbed readily through the skin. Extreme care must be exercised when preparing such gels, both prior to and after polymerization. A mask and gloves should be worn when handling the acrylamide in both solid and liquid form. Never make the assumption that the polymerized form no longer poses a potential health hazard. If you are not familiar with the mechanics of the safe handling and preparation of polyacrylamide gels, seek the assistance of someone who has experience with these techniques. Happily, it is the policy of most laboratories to purchase pre-cast polyacrylamide gels and dispense with the toxic precursors. Precast gels should also be handled with gloves.

Polyacrylamide gels¹ form as a consequence of the polymerization of acrylamide monomers into linear chains and the accompanying cross-linking of these chains with *N,N'*-methylenebisacrylamide (bis). The molecular sieving of nucleic acid molecules in a polyacrylamide gel is a direct function of the concentration of acrylamide as well as the ratio of acrylamide to bis in the gel.

Because oxygen inhibits polymerization, the monomeric mixture must be de-aerated. This may be done by purging the mixture with an inert gas or by evacuating the mixture with a vacuum.

When the gel is poured, either in a tube or as a slab, the top of the gel forms a meniscus. If ignored, this curved top on the gel will cause a distortion of the banding pattern. To eliminate the meniscus, a thin layer of water or water-saturated *n*-butanol is carefully floated in the surface of the gel mixture before it polymerizes. After polymerization, the water or butanol layer is poured off, leaving a flat upper surface on the gel. The layer of water, or water-saturated butanol, not only eliminates the meniscus but also excludes oxygen that would inhibit polymerization on the gel surface.

The pore size in an acrylamide gel may be predetermined in either of two ways. One way is to adjust the total percentage of acrylamide, that is the sum of the weights of the acrylamide monomer and the crosslinker. This is expressed as %T. For example, a 20%T gel would contain 20% (w/v) of acrylamide plus bis. As the %T increases, the pore size decreases.

¹Adapted from and courtesy of Hoefer Pharmacia Biotech.

The other way to adjust pore size is to vary the amount of crosslinker expressed as a percent of the sum of the monomer and crosslinker, or %C. For example, a 20%T5%C_{bis} gel would have 20% w/v of acrylamide plus bis, and the bis would account for 5% of the total weight of the acrylamide. It has been found that at any single %T, 5% crosslinking creates the smallest pores in a gel. Above and below 5%, the pore size increases. When there is a wide range in the molecular weights of the material under study, the investigator may prepare a pore gradient gel. The pore size in a gradient gel is larger at the top of the gel than at the bottom; the gel becomes more restrictive as the run progresses.

Polymerization of a polyacrylamide gel is accomplished by either a chemical or a photochemical method. In the most common chemical method, ammonium persulfate and the quaternary amine *N,N,N',N'*-tetramethylethylenediamine (TEMED) are used as the initiator and the catalyst, respectively. In photochemical polymerization, riboflavin and TEMED are used. The photochemical reaction is started by shining a long-wavelength ultraviolet light, usually from a fluorescent light source, on the gel mixture. Use of photochemical polymerization is limited to the study of proteins that are sensitive to ammonium persulfate or the by-products of chemical polymerization. This approach to polymerization is not suitable for the evaluation of RNA.

Polymerization generates heat. If too much heat is generated, convections will form in the unpolymerized gel, resulting in inconsistencies in the gel structure. To prevent excessive heating, the concentration of initiator-catalyst chemicals should be adjusted to complete polymerization in 20 to 60 min.

Analytical

The two denaturants that are routinely used for the electrophoresis of RNA through polyacrylamide gels are formamide (final concentration 98%; added to the loading buffer) and urea (final concentration 7M; added directly to the gel), the choice of which is mainly dependent on the size range of molecules in the sample. For molecules smaller than 200 nucleotides, urea-containing gels will prove most useful. The use of formamide-containing gels is necessary for the accurate sizing of molecules between 200 and 1000 nucleotides in length.

Slab polyacrylamide gels are poured between two clean glass plates separated by appropriate spacers (1 to 2 mm thick). When nucleic acids are electrophoresed in a vertical configuration, the gel is cooled on both sides, a dividend that generally precludes the formation "smiling bands" due to heat distortion. When recovery from the gel is required following electrophoresis, gels as thin as 0.5 mm are preferred. As with agarose gels, the most common analytical procedure involving polyacrylamide gels is the visualization of the sample after electrophoresis. In one approach, staining the gel with ethidium bromide, acridine orange, or one of the SYBR dyes permits visual inspection

of the distribution of electrophoresed nucleic acid sample. Another approach is possible when radiolabeled nucleic acids are electrophoresed. Molecules can be localized within the gel by autoradiography. Unlike standard blotting analysis (Northern blotting, for example), autoradiography under these conditions does not require transfer from the gel because either labeling and/or hybridization is conducted prior to running the gel, as in the S1 nuclease and RNase protection assays. For a complete discussion of the mechanics of polyacrylamide gel preparation and polyacrylamide electrophoresis, the reader is referred to Berger and Kimmel (1987) and Sambrook and Russell (2001).

References

- Berger, S.L. and A.R. Kimmel. (1987). *Guide to Molecular Cloning Techniques*. Academic Press, San Diego, CA.
- Sambrook, J. and D.W. Russell. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

This page intentionally left blank

Appendix N: Dot Blot Analysis

The isolation of high-quality RNA from tissue culture cells and tissue samples is merely the first (although the most critical) step in the evaluation of any model for the study of gene expression. The subsequent analysis of RNA by Northern blot analysis, nuclease protection analysis, cDNA library construction, RT-PCR, and so forth can be a time-consuming and expensive road to travel. When evaluating a new model system, cell, tissue type, or experimental regimen for the first time, it may be worthwhile to assess the system for mRNAs of interest by dot blot analysis. This simple technique allows the investigator to make “quick and dirty” statements about the biochemical composition of a sample. In addition, the dot blot approach may be useful for simply showing that the purified RNA, genomic DNA, or even cDNA is capable of hybridizing to something (anything), and that the sample warrants further characterization.

Dot blots, and the closely related variant known as slot blots, permit rapid detection of the relative amount of a particular RNA sequence both in purified samples of RNA and in cell lysates. Salient information can be obtained without electrophoresis of the sample. It is very important to understand from the onset that dot blots provide an *estimate* of the amount of target in a sample. More sophisticated methods for analysis, including cDNA synthesis or amplification mediated by the polymerase chain reaction will undoubtedly factor into the experimental design at a later stage.

In dot-blotting, RNA samples are applied directly onto a membrane under vacuum through a multi-well dot blot (Fig. N.1) or slot blot (Fig. N.2) filtration manifold, after which the samples are immobilized onto the surface of the filter membrane. Dilutions of the samples are arranged either vertically or horizontally (Fig. N.3); samples arranged in this geometry are quite simple to quantify by image analysis software (Chapter 10). The degree of hybridization is assessed by measurement of each “dot” or “slot” signal from the filter, as recorded on X-ray film. Succinctly, dot blotting is a hybridization technique without the fanfare associated with running a gel.

All types of nucleic acid molecule can be dot blotted, including poly(A)⁺ RNA, total cellular- or total cytoplasmic RNA, genomic DNA, cDNA, PCR products, and even oligonucleotides. Dot blotting is routinely used these days for screening and identification of cloned, differentially expressed genes which, for example, might have been isolated by any of a number of types of differential hybridization. When dilutions are needed, especially for titration purposes, the geometry of the dot blotter is ideal. The dot blot apparatus is easily cleaned and, if handled with care, will last for years.



Figure N.1 Minifold® I dot blot apparatus. Sample dilutions are applied under vacuum directly to the surface of the filter membrane. Depending on the format, the sample is concentrated into either 12.5 mm² “dots” or 2 mm² “spots”. Courtesy of Schleicher and Schuell Bioscience.

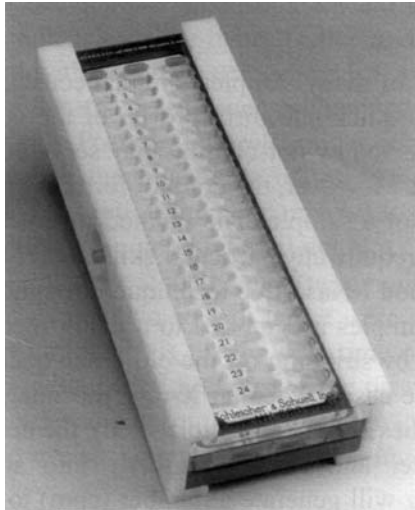


Figure N.2 Minifold® II slot blot apparatus. Sample dilutions are applied under vacuum directly to the surface of the filter membrane resting beneath the faceplate and concentrated into 6 mm² “slots”. Courtesy of Schleicher and Schuell Bioscience.

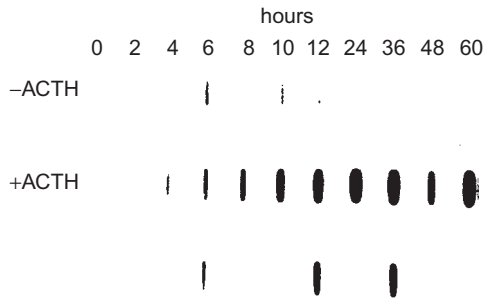


Figure N.3 Slot-blot autoradiogram of the temporal pattern of ACTH-mediated increases in P-450_{SCC} mRNA in bovine adrenocortical cells. Cells were harvested and the RNA extracted at the intervals indicated. Twenty micrograms of RNA was applied per sample to nitrocellulose using the Manifold® II slot-blotter (Schleicher & Schuell) and hybridized to a nick-translated pBSCC-2 insert. The three lower samples are duplicates of those above them. Photograph courtesy of Dr. Maliyakal John. Source: John, M.E., *et al.* (1984). Identification and characterization of cDNA clones specific for cholesterol side-chain cleavage cytochrome P-450. *Proc. Natl. Acad. Sci.* 81, 5628.

Advantages and disadvantages

Among the most obvious advantages of the dot-blotting technique is the speed with which samples are prepped for nucleic acid hybridization: a denatured sample is applied directly to a membrane and there is no gel to run. Both previously purified RNA (from any extraction method) and partially purified cell lysates, the latter of which are occasionally referred to as cytodots or quick blots (Costanzi and Gillespie, 1987), can be utilized successfully with this type of assay. Moreover, the dot blot format is ideal for making one or more dilutions of a sample. The method is rapid and facilitates the handling of many samples simultaneously. It may be used to generate a quick and dirty profile of the cellular biochemistry in one or more samples; the dot blot format is also very useful for monitoring changes in gene expression as a function of time or experimental manipulation. One might also envision the dot blot format as an ideal method for preparing in-lab panels of RNA from various tissues, different organisms, or entirely different animal or plant species. That which is observed by dot-blotting may serve as a basis for the next set of experiments.

The multi-well face plate that is the heart of the dot blot/slot blot instrument is ideally suited for the rapid preparation of replica filters. The method of dot blotting is very easy and does not require any particular skills, other than the requirement for maintaining good RNase-free technique throughout the protocol. Assuming that the samples are applied to a nylon membrane, they can be probed repeatedly by stripping off the old probe once the data have been generated. In some cases, when an X-ray film image is not satisfactory (and when the probes have been radiolabeled), the dots themselves can be cut from the filter

and dropped directly into scintillation cocktail; this will generate a number (cpm) to complement the visual data generated by autoradiography.

Of course, one can also design what might best be thought of as a reverse dot blot, in which a number of unlabeled (cold) probes are dot-blotted onto a filter, and then hybridized with labeled RNA produced, for example, by the nuclear runoff assay. This allows the investigator to screen for the expression of multiple genes simultaneously. This screening approach is described in detail in Chapter 16.

The dot blot format is an excellent way to geometrically array data. In this lab, the dot blot format has been used to follow changes in gene expression across multiple time points. This same format is also ideal for various confirmation-type assays, in which the end point is either a yes-or-no answer: for example, either a sample hybridizes to a probe or it doesn't. The format lends itself to very easy scanning using any of the standard image analysis software packages.

The disadvantages of performing dot blots are closely related to the mechanics of how the method is performed. As there is no gel involved, the sample being applied directly to a membrane, dot-blotting renders quantitative data only; signal strength correlates with the abundance of a transcript and there is no way to determine the molecular weight of the transcript(s). Further, using this method alone, one cannot discern how many different sized transcripts have managed to hybridize to the probe. This is especially problematic when the probe manifests even a minimal affinity for the enormously abundant ribosomal RNAs. The potential for high background hybridization thus mandates the inclusion of excellent positive and negative controls. As with all nucleic acid assays in which a nylon, nitrocellulose, or polyvinylidene difluoride (PVDF) membrane is involved, the quantitiveness of this assay is compromised because the act of physical immobilization onto the surface of a filter membrane renders some of the molecules incapable of nucleic acid hybridization.

Appropriate positive and negative controls

Dot blot and slot blot analyses yield quantitative data which is semi-, semi-quantitative at best. The main drawback of this approach is that it lacks the qualitative component that accompanies electrophoresis. To be truly reliable, dot blot analysis must include excellent positive and negative controls, in order to demonstrate hybridization specificity and to gauge non-specific probe binding to the filter membrane. For example, applying dilutions of rRNA or tRNA to unused wells would show the degree of cross-hybridization to the non-poly(A)⁺ component of the sample, especially the ribosomal RNAs. Further, application of the popular λ -*HindIII* digested- and Φ X174-*HaeIII*-digested bacteriophage genomes (used as DNA molecular weight standards on agarose

gels) should yield no observable hybridization to probe molecules if the hybridization and post-hybridization washes are conducted with adequate stringency. If these popular molecular weight DNA standards or other RNA standards are not available, *any* DNA that is unrelated to the probe will do, including previously characterized PCR products that might otherwise be discarded. It may also be useful to apply nothing but buffer to at least one well to demonstrate lack of buffer-associated signal. Positive controls might include dilutions of cDNA complementary to the transcripts of interest (the probe itself), which, depending on the dilution prepared, may well yield the strongest signal on the filter.¹ Moreover, good internal dot blot controls are always in order: one should observe equally intense signals from wells into which equal amounts of positive control target were applied. When attempting this type of blot analysis for the first time or with a new system, it is strongly suggested that dilutions of the positive control target material be made in order to determine the linear range of the assay. For example, it would be useless, quantitatively speaking, if the hybridization signals were too intense to be accurately measured on X-ray film (recall that all films exhibit a rather narrow linear range).

When choosing between dot blots or slot blots, it may be useful to consider the area into which samples are concentrated by the required manifold. The slot configuration often generates a higher signal to noise ratio because of the smaller surface area into which the sample is concentrated. (2 mm² for spots, 6 mm² for slots, and 12.5 mm² for dots). Thus, the more quantitative slot blot configuration may be helpful when working with low mass quantities or in the assay of very low abundance transcripts. Practically speaking, however, it really does not make a great deal of difference how the sample is configured because of the rather low sensitivity of this assay. Dot blot data are often used as a stepping stone to screening techniques associated with a higher level of quantitiveness.

Protocol²: RNA dot blots

1. *Wear gloves!* Be sure that all reagents are purged of nuclease activity prior to contact with RNA samples.
2. Purify RNA according to any of the protocols presented in this volume or elsewhere. RNA should be dissolved in sterile H₂O or TE buffer, pH 7.5 (10 mM Tris-Cl, 1 mM EDTA, pH 7.5). Plan to apply 1–10 µg of RNA per dot/slot in a volume of 100–200 µl per well. Dilute only RNA that will be used for this assay.

¹Because of the very strong signals usually generated by positive controls, it is best to prepare a substantial dilution of the positive control material before applying it to the filter. Moreover, because strong signals from dot blots tend to bleed over and obscure signals from proximal wells, it is also best to avoid positioning experimental samples in those wells adjacent to the positive control.

²Protocol adapted, in part, from Schleicher & Schuell (1995).

As always, stock RNA should be stored as concentrated as possible at -80° for future experimentation.

3. For each 100 μ l of diluted RNA, add 60 μ l 20 \times SSC buffer and 40 μ l 37% formaldehyde stock solution³ to the RNA sample. Incubate at 60° for 15 min to denature the sample.

Note: The manifold should be assembled at this time so that the sample may be applied to the membrane immediately following denaturation.

4. While the RNA is denaturing, pre-wet the nylon or other filter in autoclaved water for 5 min. Be sure to handle the filter as little as possible and do so only with gloves. Equilibrate the filter in 6 \times SSC just prior to use and saturate two sheets of absorbent blotting paper (Schleicher & Schuell #GB003, or the equivalent) in 6 \times SSC as well.

Note: The ionic strength required is dependent on the chemical composition and surface charge of the filter. As always, it is best to follow the manufacturer's instructions for equilibrating any filter prior to any application.

5. Place the saturated sheets of GB003 blotting paper on the filter support plate of the filtration manifold. Place the nylon filter on top of the blotting paper, and clamp the sample faceplate into position. When correctly assembled, the O-rings are in direct contact with the filter and the wells are facing up.
6. Apply a low vacuum to the dot blot device and ensure that the residual buffer from the pre-wetting step is being drawn through the membrane. Wash individual wells with 500 μ l of 6 \times SSC.
7. Apply samples into the wells in a volume of 100 to 400 μ l per well. Dilutions, if required, are made in 6 \times SSC.
8. After all of the sample has been pulled through the nylon filter, wash each well with an additional 300 μ l of 6 \times SSC. When this wash aliquot has been pulled through the membrane, disconnect the vacuum source and remove the membrane from the manifold.
9. Immobilize the RNA onto the filter membrane according to the instructions provided by the manufacturer, usually by UV crosslinking (see Chapter 11 for immobilization strategies). If the filter will not be probed right away, store it in a cool, dry place, out of direct light.
10. Perform prehybridization blocking, probe hybridization, and posthybridization stringency washes (parameters described in Chapter 13).
11. Perform detection protocols appropriate for the method by which the probe was labeled (discussed in Chapter 14).

Protocol⁴: DNA dot blots

The ability to perform rapid nucleic acid analysis by dot-blotting also extends into the realm of DNA characterization. DNA dot blotting is identical to the methods prescribed for RNA dot blots, the only exception being the method

³Formaldehyde (HCHO) must be freshly deionized prior to use to remove formic acid and, in so doing, drive the pH above 4.0. Formamide (HCONH₂) must also be deionized prior to use. Formaldehyde and formamide are both toxic and should be handled with care according to the manufacturer's MSDS.

⁴Protocol adapted, in part, from Schleicher & Schuell (1995).

for double-stranded DNA denaturation. The same concerns addressed above for RNA also apply to the design of DNA dot blots.

1. *Wear gloves!* Be sure that all reagents are purged of nuclease activity prior to contact with DNA samples.
2. Purify DNA according to any of the widely available standard protocols for genomic DNA isolation or cDNA synthesis. A convenient genomic DNA stock concentration for most routine assays is 1.0mg/ml. DNA should be dissolved in TE buffer, pH 8.0 (10mM Tris-Cl, 1mM EDTA, pH 8.0). Plan to apply 5 to 10 μ g of DNA per dot/slot in a volume of 100 to 200 μ l per well, and dilute the DNA sample as needed in TE buffer or sterile H₂O.

Note: It may be necessary to add even greater amounts of DNA per well to be able to detect single copy sequences.

3. For each 100 μ l DNA in TE buffer, add 10 μ l 2N NaOH and incubate at 37° for 10 min. Add 40 μ l 20 \times SSC buffer and then place on ice if not prepared to apply to the membrane immediately. As an alternative to the addition of 20 \times SSC, one may add an equal volume of 2M NH₄OAc, pH 7.0, and likewise place the sample on ice.

Note 1: The manifold should be assembled at this time so that the sample may be applied to the membrane immediately following denaturation.

Note 2: It is always best to apply denatured samples to the filter as soon as possible so as to preclude possible reannealing of the denatured, single-stranded molecules.

4. While the DNA is denaturing, pre-wet the nylon or other filter in autoclaved water for 5 min. Be sure to handle the filter as little as possible and do so only with gloves. Equilibrate the filter in 6 \times SSC just prior to use, and saturate two sheets of absorbent blotting paper (Schleicher & Schuell #GB003, or the equivalent) in 6 \times SSC as well.

Note: The ionic strength required is dependent on the chemical composition and surface charge of the filter. As always, it is best to follow the manufacturer's instructions for equilibrating any filter prior to any application.

5. Place the saturated sheets of GB003 blotting paper on the filter support plate of the filtration manifold. Place the nylon filter on top of the blotting paper, and clamp the sample faceplate into position. When correctly assembled, the O-rings are in direct contact with the filter and the wells are facing up.
6. Apply a low vacuum to the dot blot device and ensure that the residual buffer from the pre-wetting step is being drawn through the membrane. Wash individual wells with 500 μ l of 6 \times SSC.
7. Apply samples to the wells in a volume of 100 to 400 μ l per well. Dilutions, if required, are made in 6 \times SSC.
8. When the entire sample has been pulled through the nylon filter, wash each well with an additional 300 μ l of 6 \times SSC. After the wash aliquot has been pulled through the membrane, disconnect the vacuum source and remove the membrane from the manifold.
9. Immobilize the DNA onto the filter membrane according to the instructions provided by the manufacturer, usually by UV crosslinking (see Chapter 11 for immobilization strategies). If the filter will not be probed right away, store it in a cool, dry place, out of direct light.
10. Perform prehybridization blocking, probe hybridization, and posthybridization stringency washes (parameters described in Chapter 13).
11. Perform detection protocols appropriate for the method by which the probe was labeled (discussed in Chapter 14).

Limitations of the data

The greatest limitations of data generated by dot-blotting are the lack of a qualitative aspect because there is no gel and, because of the mechanics of the assay, the compromised sensitivity, compared to other standard techniques for the assay of gene expression. One should also be acutely aware that highly concentrated nucleic acid samples, including those suspended in relatively high salt buffers, are notorious for clogging filters when the sample is applied to the well of the dot blot manifold. Further, overloading the wells can result in “spot saturation”, meaning that the investigator has applied more material than the filter can bind, or has diluted the sample in a reagent that reduces the binding capability of the membrane.

The dot blot assay, while semiquantitative, has a relatively narrow linear range, due to the ease with which one can exceed the binding capacity of the filter as well as the narrow linear range of X-ray films. Samples that generate enormous signals will literally burn out the film so that the true differences between the dots are completely obscured. Never forget that image analysis software can only analyze that which the user provides, and it is incumbent upon the investigator to at least try to minimize ambiguity wherever possible. For these reasons, dilutions of each sample are strongly recommended, at least when performing this assay for the first time with uncharacterized material, in order to assess the dynamic range of the sample. Following autoradiography or detection by chemiluminescence, the individual dots may be cut out and the cpm measured if a number is needed to go along with the picture. While this additional numerical data may be quite helpful in this regard, it should also be obvious that cut up filters cannot be used again. As suggested above, there really is no method by which to assess cross hybridization within a particular spot on the filter, though well-thought-out negative and positive controls can provide fairly convincing, though indirect, evidence of the fidelity of the assay.

Lastly, because of the relative ease with which investigators can make cDNA, perform RT-PCR, run gels, and perform nucleic acid blot hybridization, dot blotting may not be the best choice at the onset of study. Dot blotting is a “graphical” assay: a panel of dots representing hybridized samples is nice to look at, but does little more than organize the data. If an assay with geometric format is desired, then blotting and probing the samples is in order. For more quantitative approaches, the investigator should select an alternative technique.

References

- Costanzi, C. and D. Gillespie. (1987). Fast blots: immobilization of DNA and RNA from cells. In “*Guide to Molecular Cloning Techniques*” (S.L. Berger and A.R. Kimmel, Eds). Academic Press, San Diego, CA. pp. 582–587.
- Schleicher & Schuell, Inc. (1995). Nucleic acid dot/slot-blots onto S&S transfer media. Nucleic acid dot/slot-blots onto S&S transfer media. In “*Blotting, Hybridization, and Detection*,” 6th edition. Schleicher & Schuell. Keene, NH, pp. 20–22.

Appendix 0: Centrifugation as a Mainstream Tool for the Molecular Biologist

Centrifugation is a separation technique based on the fact that objects moving in a circular path are subjected to an outward-directed force. The magnitude of this force, commonly expressed in terms of the earth's gravitational force (relative centrifugal force [RCF] or the “number times g ”) is a direct function of the radius of rotation and the angular velocity. Centrifugation is an indispensable procedure for the separation of whole cells, organelles, and macromolecules (also referred to as particles) from a solution, based on size and density. As such these instruments are in constant use in most laboratories. Improper handling and maintenance, however, can cause permanent injury or worse. It is the responsibility of the user, especially in the case of ultracentrifuges, to ensure that the centrifuge instrument and the centrifuge rotor are operated within the design limits of the machine. This includes derating the rotor, as described in the manual that accompanied its purchase. For excellent overviews of centrifugation, see Cooper (1977) and Greene and Castora (2009).

Types of centrifuges

Desk top clinical centrifuges

- Usually operate below 3000 rpm and at ambient temperature.

High speed centrifuges

- Operate between 20,000–25,000 rpm and are usually refrigerated.

Desk top microcentrifuges (microfuges)

- Operate up to 14,000 rpm ($12,000 \times g$).
- Are sometimes equipped with refrigeration.
- Are indispensable for nucleic acid samples in small volumes.

Ultracentrifuges

- Operate up to $500,000 \times g$ (75,000 rpm with $r = 8$ cm).
- Permit fractionation of subcellular organelles.

- Permit fractionation of molecules based strictly on density.
- Have a wide temperature range available.
- Are vacuum-operated (air friction severe above 40,000 rpm).

Rotors

Centrifuges are designed to accommodate the specific requirements for a given separation. Parameters include temperature, RCF, volume of sample, duration of run, shape of gradient (e.g., linear or step), choice of differential or density gradient separation, and type of rotor. A fixed angle rotor (Fig. O.1) is one in which the sample is maintained at a defined angle during the centrifugation period. In contrast, swinging bucket rotors (Fig. O.2) allow the holder or bucket

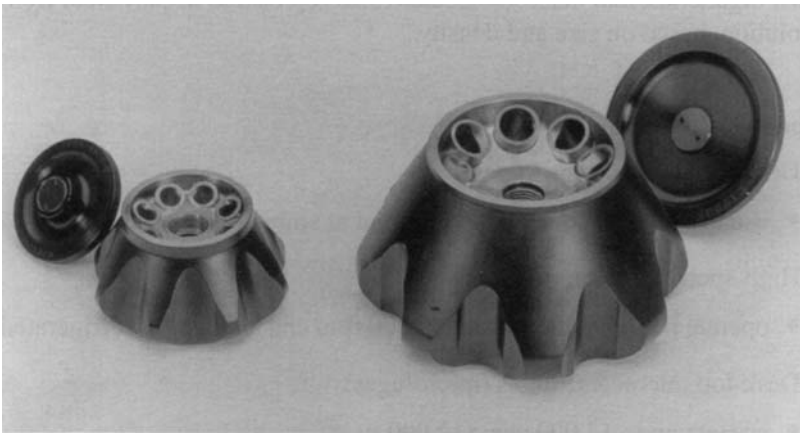


Figure O.1 Fixed angle rotor. Samples are held in fixed position during the centrifugation. Courtesy of Sorvall, Inc.



Figure O.2 Swinging bucket rotor. Buckets holding the samples swing outward 90 degrees with respect to the axis of rotation. Courtesy of Sorvall, Inc.

into which the sample is placed to swing outward to a position 90° , with respect to the axis of rotation. Vertical rotors maintain sample tubes upright throughout the run, parallel to the axis of rotation. The choice of fixed angle, swinging bucket, or vertical rotor is completely dependent on the intended application. Fixed angle rotors are most compatible with differential centrifugation techniques, while efficient gradient-based separations are supported by swinging bucket and vertical rotors.

A useful method by which to correlate performance among rotors is comparison of their respective k or k' factors, values used to compare efficiency among rotors in a particular application. The k factor estimates the time required to pellet a particle, when the sedimentation coefficient of the particle is known; the k' factor is indicative of the time required to move a zone of particles to the bottom of a centrifuge tube. Succinctly, the lower the k factor, the more efficient the rotor. The characteristics of each rotor are unique and the manufacturer's specifications are available on-line for comparative shopping.

Applications¹

In the molecular biology laboratory, centrifugation fractionation techniques are utilized at several levels during the isolation of nucleic acids, including the pelleting of whole cells prior to cell lysis, phase separation during nucleic acid extraction, gradient formation during isopycnic separation, and collection of salt and alcohol-precipitated material. Be sure to use only those high-strength tubes that can withstand the required RCF. In the case of consumable plasticware, this information accompanies the product when it is shipped. Corex glass tubes, referred to in several protocols in this book, are often used at RCFs greater than can be tolerated by polypropylene products. If Corex tubes are not available, an equivalent product can be purchased from Kimble-Kontes (use catalogue number 45500-15 for 15 ml tubes and 45500-30 for 30 ml tubes). As with Corex tubes, the Kimble high-strength tubes must be placed in a suitable rubber adapter for centrifugation. Ultracentrifugation, however, requires tubes that are designed specifically for that application. Table O.1 is intended to be a general guide only; the precise RCF maxima, as well as solvent compatibility, for each product are described by the manufacturer, and those specifications should be adhered to tenaciously. Moreover, the maximum RCF may be increased, in some cases, by placing tubes in adapters within the rotor. It is also critical to note that exposing centrifuge tubes to chemicals with which the tubes are incompatible is a recipe for tube failure, not to mention possible danger to lab personnel and damage to the centrifuge and rotor. Always check for chemical compatibility before loading or treating centrifuge tubes in any way.

¹Adapted, in part, from Griffith, 1986.

Table O.1 Selected Centrifuge Tubes for Common Molecular Biology Applications

Type of plastic	Tube size	Typical ² maximum RCF	Common centrifuge(s)
Polypropylene	1.5 ml microfuge	15,000 × <i>g</i>	Microcentrifuge, Clinical centrifuge
	15 ml conical	3,000 × <i>g</i>	
	50 ml conical	5,000 × <i>g</i>	
Polystyrene	15 ml conical	1,500 × <i>g</i>	Clinical centrifuge
	50 ml conical	1,800 × <i>g</i>	Clinical centrifuge
Corex glass	15 ml	12,000 × <i>g</i>	Clinical centrifuge, high-speed centrifuge
	30 ml	12,000 × <i>g</i>	
Oak Ridge Tubes (polycarbonate)	10–50 ml sizes	50,000 × <i>g</i> (tube must be 80% full)	High-speed centrifuge
Polyallomer	Thin wall	100,000–	Ultracentrifuge
		500,000 × <i>g</i>	
	Thick wall	70,000– 175,000 × <i>g</i>	

Differential centrifugation

Differential centrifugation is the simplest, most straightforward centrifugation technique for sample fractionation. In this method, centrifuge tubes containing a homogeneous sample mixture are subjected to a brief centrifugation (usually less than 30 min). At the conclusion of the run, the pellet at the bottom of the tube includes all material sedimented during the run. It should be clear that the pellet is also contaminated with anything that was at or close to the bottom of the tube from the onset. The supernatant, containing non-sedimented material, is usually removed by decanting; in the case of a very firm pellet, the supernatant can be removed carefully by aspiration. Should further fractionation be required, centrifugation of the supernatant in a fresh tube for a longer period than the initial centrifugation, or at higher speeds, will produce a new pellet and supernatant.

Density gradient centrifugation–sedimentation velocity

A more sophisticated preparative technique is density gradient centrifugation, in which sample particles move through a density gradient to achieve separation.

²Maximum tolerable RCF varies widely by tube design, tube size, manufacturer, and actual volume within the tube. This table is intended to be a general guide only. Be sure to check the specifications for the specific tubes in use in the lab.

The density of the gradient is lowest at the top of the tube and increases toward the bottom of the tube. In one type of density gradient centrifugation, known as sedimentation velocity, or rate zonal centrifugation, a sample is layered on top of a preformed gradient. When subjected to a centrifugal force, particles in the sample begin to move downward through the gradient in discrete zones; the rate of zone movement is governed by the sedimentation rate of individual particles within a zone. The hallmark of this type of separation is centrifugation through a relatively shallow gradient at low speeds for a short time (compared to buoyant density centrifugation). Successful sedimentation velocity centrifugation requires the following conditions:

1. The density of particles in the sample must be greater than the density of the gradient at every point throughout the gradient.
2. Centrifugation must be terminated before any of the separated material (specifically the zone of greatest density) reaches the bottom of the tube.

Example: In one application, a heterogeneous sample of RNA can be size-fractionated by centrifugation through a sucrose gradient (Benecke *et al.*, 1978; Nevins and Darnell, 1978). A typical sucrose gradient is as shallow as 5 to 20% to as steep as 5 to 40%; in the case of the higher viscosity 5 to 40% gradient, resolution can be improved by increasing the centrifugation speed. According to the parameters defined here, RNA molecules would move through the gradient according to their sedimentation rate. At the conclusion of the run, the RNA will be distributed throughout the gradient, based on size. In one type of enrichment, fractions of the chromatographed RNA are removed from the gradient for further characterization.

Density gradient centrifugation – isopycnic technique

Another type of separation through a density gradient is known as isopycnic, buoyant density, or density equilibrium centrifugation. In this application, particles in a sample move through a gradient only to the point at which the density of the gradient is equal to the density of the particle and at which point particles of identical density float or band. Unlike sedimentation velocity centrifugation, extending the period of centrifugation will not result in the continued downward migration of the sample through the gradient. The gradient itself must be very steep, and at its greatest point it must exceed the density of the particles of interest. Note that very dense components of a sample may become pelleted while other components migrate only to their isopycnic positions. Commonly used materials for this type of separation include CsCl, Cs₂SO₄, and CsTFA. Cesium chloride, historically the salt used for the classical studies of semiconservative replication (Meselson *et al.*, 1957; Meselson and Stahl, 1958), is routinely used to establish gradients with densities ranging up to about 1.8 g/cm³. Cs₂SO₄ and CsTFA can be used to form a gradient twice as steep as that which

is achievable with the chloride salt and the former of which is preferred for separating DNAs with widely different buoyant densities.

A gradient for buoyant density centrifugation need not be preformed. For example, in several applications, solid CsCl is added to a nucleic acid mixture (chromosomal DNA, RNA, protein, plasmid DNA). Under the centrifugal force experienced during ultracentrifugation, the gradient material redistributes, due in part to the intrinsic density of cesium salts. These self-forming gradients traditionally require several hours at ultracentrifugation speeds before macromolecules become isopycally banded, though the required centrifugation time can be significantly shortened by using a micro-ultracentrifuge. In the resulting linear gradient generated during the run, components of the sample will either sediment or float to their isopycnic locations, based only on their respective densities.

References

- Benecke, B.J., A. Ben-Ze'ev, and S. Penman. (1978). The control of mRNA production, translation, and turnover in suspended and reattached anchorage-dependent fibroblasts. *Cell* **14**, 931.
- Cooper, T.G. (1977). *The Tools of Biochemistry*. John Wiley & Sons, New York pp. 309–351.
- Greene, J.J. and F. Castora (Eds). (2009). *Handbook of Biologic Research Methodology*. CRC Press.
- Griffith, O.M. (1986). *Techniques of Preparative, Zonal, and Continuous Flow Ultracentrifugation*. Beckman Instruments, Inc., Palo Alto, CA.
- Meselson, M. and F.W. Stahl. (1958). The replication of DNA in *E. coli*. *Proc. Natl. Acad. Sci.* **44**, 671.
- Meselson, M., F.W. Stahl, and J. Vinograd. (1957). Equilibrium sedimentation of macromolecules in density gradients. *Proc. Natl. Acad. Sci.* **43**, 581.
- Nevins, J.R. and J.E. Darnell. (1978). Steps in the processing of Ad2 mRNA: Poly (A)⁺ nuclear sequences are conserved and poly(A)⁺ addition precedes splicing. *Cell* **15**, 1477.

Appendix P: Selected Suppliers of Equipment, Reagents, and Services

AirClean Systems

3248 Lake Woodard Drive
Raleigh, NC 27604
800/849-0472
www.aircleansystems.com

Ambion, Inc.

2130 Woodward
Austin, TX 78744
800/888-8804
www.ambion.com

Amersham/GE Healthcare Life Sciences

800 Centennial Avenue
P.O. Box 1327
Piscataway, NJ 08855
800/526-3593
www.gelifesciences.com

Applied Biosystems

850 Lincoln Centre Drive
Foster City, CA 94404
800/345-5224
www.appliedbiosystems.com

Aptagen

250 North Main Street
Jacobus, PA 17407
877/278.2463
www.aptagen.com

BD Biosciences

2350 Qume Drive
San Jose, CA 95131
877/232-8995
www.bdbiosciences.com

Beckman Coulter, Inc.

4300 N. Harbor Boulevard
Box 3100
Fullerton, CA 92834
800/742-2345
www.beckmancoulter.com

Bellco Glass, Inc.

340 Edrudo Road
Vineland, NJ 08360
800/257-7043
www.belcoglass.com

Bio-Rad Laboratories

1000 Alfred Nobel Drive
Hercules, CA 94547
800/424-6723
www.biorad.com

BioSpec Products, Inc.

P.O. Box 788
Bartlesville, OK 74005
800/617-3363
www.biospec.com

Carestream Molecular Imaging

4 Science Park
New Haven, CT 06511
877/747-4357
www.carestreamheath.com

Clontech Laboratories, Inc.

1290 Terra Bella Avenue
Mountain View, CA 94043
800/662-2566
www.clontech.com

Cole-Parmer Instrument Company

625 E. Bunker Court
Vernon Hill, IL 60021
800/323-4340
www.coleparmer.com

Dharmacon

2650 Crescent Drive, #100
Lafayette, CO 80026
800/235-9880
www.dharmacon.com

Diversified Biotech, Inc.

1208 V.F.W. Parkway
Boston, MA 02132
800/796-9199
www.divbio.com

DNA Diagnostics, Inc.

P.O. Box 4544
Crofton, MD 21114
410/286-0092
www.DNAdiagnosticsinc.com

Eastman Kodak Co.

343 State Street
Rochester, NY 14650
800/242-2424
www.kodak.com

Exon-Intron, Inc.

Specialists in Biotech Education
P.O. Box 395
Loganville, PA 17342
717/428-0270
www.DNAtech.com

FOTODYNE, Inc.

950 Walnut Ridge Drive
Hartland, WI 53029
800-362-3686
www.fotodyne.com

GenHunter Corporation

624 Grassmere Park Drive
Suite 17
Nashville, TN 37211
800-311-8260
www.genhunter.com

Hoefer, Inc.

84 October Hill Road
Holliston, MA 01746
800/227-4750
www.hoeferinc.com

Invitrogen Corporation

1600 Faraday Avenue
Carlsbad, CA 92008
800/955-6288
www.invitrogen.com

JustRight Technology

821 Frederick Road, Suite 201
Catonsville, MD 21228
410/747-2001
www.justright.com

Kimble Chase Life Science

1022 Spruce Street
Vineland, NJ 08362
888-546-2531
www.kimble-kontes.com

Kinematica, Inc.

1648 Locust Avenue
Unit C
Bohemia, NY 11716
631-750-6653
www.kinematica-inc.com

Kirkegaard & Perry Labs., Inc.

910 Clopper Road
Gaithersburg, MD 20878
800/638-3167
www.kpl.com

Lofstrand Labs, Ltd.

7961 Cessna Avenue
Gaithersburg, MD 20879
800/541-0362
www.lofstrand.com

Lonza Rockland, Inc.

191 Thomaston Street
Rockland, ME 04841
800/341-1574
www.lonza.com

Lumigen, Inc.

22900 W. Eight Mile Road
Southfield, MI 48033
248/351-5600
www.lumigen.com

Matrix Technologies Corporation

22 Friars Drive
Hudson, NH 03051
800/345-0206
www.matrixtechcorp.com

Media Cybernetics, Inc.

4340 East-West Highway
Suite 400
Bethesda, MD 20814
800/992-4256
www.mediacy.com

Millipore Corporation

290 Concord Road
Billerica, MA 01821
800/645-5476
www.millipore.com

MO BIO Laboratories, Inc.

2746 Loker Avenue West
Carlsbad, CA 92010
800/606-6246
www.mobio.com

MP Biomedicals

29525 Fountain Parkway
Solon, OH 44139
800/854-0530
www.mpbio.com

Nalge Nunc International

75 Panorama Creek Drive
Rochester, NY 14625
800/625-4327
www.nalgenunc.com

NanoDrop Technologies

3411 Silverside Rd
Bancroft Building
Wilmington, DE 19810
302/479-7707
www.nanodrop.com

New England Biolabs

240 County Road
Ipswich, MA 01938
800/632-5227
www.neb.com

Nikon, Inc.

1300 Walt Whitman Road
Melville, NY 11747
800/526-4566
www.nikonusa.com

Omni International

1000 Williams Drive
Suite 1024
Marietta, GA 30066
800/776-4431
www.omni-inc.com

Owl Separation Systems

75 Panorama Creek Drive
Rochester, NY 14625
800/242-5560
www.owlsoci.com

PerkinElmer Life and Analytical Sciences, Inc.

940 Winter Street
Waltham, MA 02451
800/762-4000
www.perkinelmer.com

Pierce Biotechnology

P.O. Box 117
Rockford, IL 61105
800/874-3723
www.piercenet.com

Polaroid Corporation

300 Baker Avenue
Concord, MA 01742
800/225-1618
www.polaroid.com

Promega Corporation

2800 Woods Hollow Road
Madison, WI 53711
800/356-9526
www.promega.com

Roche Applied Science

9115 Hague Road
P.O. Box 50414
Indianapolis, IN 46250
800/262-1640
www.roche-applied-science.com

Rohm & Haas Co.

100 Independence Mall West
Philadelphia, PA 19106
877/288-5881
www.rohmhaas.com

Sigma-Aldrich

3050 Spruce Street
St. Louis, MO 63160
800/325-3010
www.sigmaaldrich.com

Stratagene

11011 North Torrey Pines Road
La Jolla, CA 92037
800/424-5444
www.stratagene.com

Teledyne Isco, Inc.

4700 Superior Street
Lincoln, NE 68504
800/228-4250
www.isco.com

Transponics

250 North Main Street
Jacobus, PA 17407
717/428-0575

Tri-Continent Scientific, Inc.

12555 Loma Rica Drive
Grass Valley, CA 95945
800/937-4738
www.tricontinent.com

VIRxSYS Corporation

200 Perry Parkway, Suite 1A
Gaithersburg, MD 20877
301/987-0480
www.virxsys.com

VWR Scientific

1310 Goshen Parkway
West Chester, PA 19380
800/932-5000
www.vwrsp.com

Whatman, Inc.

Building 1
800 Centennial Avenue
Piscataway, NJ 08854
800/526-3593
www.whatman.com

Worthington Biochemical Corporation

730 Vassar Avenue

Lakewood, NJ 08701

800/445-9603

www.worthington-biochem.com**Zeiss, Inc.**

One Zeiss Drive

Thornwood, NY 10594

800/233-2343

www.zeiss.com

Appendix Q: Useful SI Units

Factor	Prefix	Symbol
10^{24}	yotta	Y
10^{21}	zetta	Z
10^{18}	exa	E
10^{15}	penta	P
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^2	hecto	h
10^1	deka	da
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f
10^{-18}	atto	a
10^{-21}	zepto	z
10^{-24}	yocto	y

This page intentionally left blank

Appendix R: Common Abbreviations

A: adenine
amol: attomole
ARE: adenine/uracil-rich element
BMP: bit-map
C: cytosine
cRNA: complementary (antisense) RNA
cpRNA: chloroplast RNA
ctRNA: countertranscript RNA
DNA: deoxyribonucleic acid
dNTP: a cocktail of nucleotides containing equimolar amounts of
dATP, dCTP, dGTP, dTTP
EDTA: ethylenediaminetetraacetic acid
fg: femtogram
G: guanine
GIF: graphics interchange format
h: hour
hnRNA: heterogeneous nuclear RNA
ISO: International Organization for Standardization
JPEG: joint photographic experts group
μg: microgram (10^{-6} gram)
μl: microliter (10^{-6} liter)
min: minutes
mol: mole
mRNA: messenger RNA
MSDS: material safety data sheet
mtRNA: mitochondrial RNA
NaOAc: sodium acetate
ng: nanogram
NTP: a cocktail of nucleotides containing equimolar amounts of
ATP, CTP, GTP, UTP
nt: nucleotide
OD: optical density
PCI: phenol:chloroform:isoamyl alcohol (25:24:1)
PCR: polymerase chain reaction
pg: picogram
Pu: purine (adenine, guanine)
Py: pyrimidine (cytosine, thymine, uracil)
q.s.: *quantum sufficiat, quantum sufficit, quantum satis*
rDNA: recombinant DNA

RNA: ribonucleic acid

RNase: ribonuclease

RPA: ribonuclease protection assay

rRNA: ribosomal RNA

s, sec: second

SAM: S-adenosyl-methionine

T: thymine

TMB: 3,3',5,5'-tetramethylbenzidine

tRNA: transfer RNA

U: uracil

1° : primary

2° : secondary

3° : tertiary

Appendix S: Trademark Citations

AccuTaq LA™ is a registered trademark of the Sigma-Aldrich Corporation.
Amerlite® is a registered trademark of Rohm and Hass.
Amplifluor® is a registered trademark of Intergen Company.
Ampligen® is a registered trademark of Hemispherx Biopharma.
AmpliTaq® is a registered trademark of Roche Molecular Biosystems, Inc.
AmpliWax® is a registered trademark of Roche Molecular Biosystems, Inc.
Aptabody™ is a registered trademark of Aptagen, LLC.
Beadbeater™ is a trademark of BioSpec, Inc.
BioMax™ is a trademark of Eastman Kodak Co.
BLAST® is a registered trademark of the National Library of Medicine.
Centricon® is a registered trademark of Millipore Corporation.
CDP Star® is a registered trademark of Applied Biosystems.
Corex® is a registered trademark of Corning, Inc.
Coomassie® is a registered trademark of Imperial Chemical Industries, Ltd.
CSPD® is a registered trademark of Applied Biosystems.
Cy3® is a registered trademark of Amersham Biosciences.
Cy5® is a registered trademark of Amersham Biosciences.
Deep Vent® is a registered trademark of New England BioLabs.
DNA Dipstick™ is a trademark of Invitrogen Corporation.
Dynabeads® is a registered trademark of Dynal Biotech, ASA.
ECL™ is a trademark of GE Healthcare.
E-Gel® is a registered trademark of Invitrogen Corporation.
ELIMINase® is a registered trademark of Decon Laboratories.
Elutip-d® is a registered trademark of Schleicher & Schuell Bioscience.
Eppendorf® is a registered trademark of Eppendorf-NetherlerHinz GmbH.
Expand™ is a trademark of Roche Applied Science.
Extractor® is a registered trademark of Scheicher & Schuell Bioscience.
Ficoll® is a registered trademark of Amersham Biosciences.
FORMAZol® is a registered trademark of Molecular Research Center, Inc.
Gel-Pro® is a registered trademark of Media Cybernetics, Inc.
GelStar® is a registered trademark of FMC Corporation.
GenBank® is a registered trademark of the United States Department of Health and Human Services.
GeneAmp® is a registered trademark of Roche Molecular Systems.
Hybond® is a registered trademark of Amersham Biosciences.
Hyperfilm™ is a trademark of Amersham Biosciences.
IGEPAL® is a registered trademark of General Dyestuff Corp.
JumpStart™ is registered trademark of the Sigma Aldrich Corporation.
Kimwipes® is a registered trademark of the Kimberly-Clark Corporation.

Lab-Line™ is a trademark of Lab-Line Instruments, Inc.
Lumiphos® is a registered trademark of Lumigen, Inc.
Minifold® is a registered trademark of Schleicher & Schuell Bioscience.
Nalgene® is a registered trademark of Nalge Nunc International Corporation.
Nonidet P-40™ is a trademark of Shell International Petroleum Company Ltd.
Nytran® is a registered trademark of Schleicher & Schuell Bioscience.
Optitran™ is a registered trademark of Whatman.
Parafilm® is a registered trademark of American Can Company.
pBluescript® is a registered trademark of Stratagene
pGEM® is a registered trademark of the Promega Corporation.
PhosphorImager® is a registered trademark of GE Healthcare Life Sciences.
Plexiglass™ is a trademark of Rohm and Haas Co.
Polaroid® is a registered trademark of the Polaroid Corporation.
Polytron® is a registered trademark of Kinematica AG Company.
PosiBlot® is a registered trademark of Stratagene.
Pronase® is a registered trademark of Calbiochem-Novabiochem Corp.
PYREX® is a registered trademark of Corning, Inc.
RedTaq™ is a trademark of the Sigma-Aldrich Corporation.
Riboprobe® is a registered trademark of Promega Corporation.
RNAGuard® is a registered trademark of Amersham Biosciences.
RNAlater® is a registered trademark of Ambion, Inc.
RNase Away® is a registered trademark of Mo Bio Laboratories.
RNase-Off™ is a trademark of Takara Mirus Bio.
RNaseZap® is a registered trademark of Ambion, Inc.
RNasin® is a registered trademark of Promega Corporation.
Rubbermaid® is a registered trademark of the Newell Rubbermaid Corporation.
SaranWrap™ is a trademark of Dow Chemical.
SeaKem GTG® is a registered trademark of Cambrex Bio Science.
Sephadex™ is a trademark of Amersham Biosciences.
Sepharose™ is a trademark of Amersham Biosciences.
Speed Vac® is a registered trademark of Savant Instruments, Inc.
Sigmacote® is a registered trademark of Sigma-Aldrich Company.
STORM® is a registered trademark of GE Healthcare Life Sciences.
Stratalinker® is a registered trademark of Stratagene.
Subtractor® is a registered trademark of Invitrogen Corporation.
SUPERase•In™ is a trademark of Ambion, Inc.
Super Script™ is a trademark of Invitrogen Corporation.
SYBR® is registered trademark of Molecular Probes, Inc.
SYPRO® is registered trademark of Molecular Probes, Inc.
TAMRA™ is a trademark of Perkin Elmer.
Triton® is a registered trademark of Union Carbide Chemicals and Plastics Co., Inc.
TaqMan® is a registered trademark of Applied Biosystems.
TaqStart™ is a trademark of Clontech Laboratories, Inc.
Teflon® is a registered trademark of E.I. DuPont de Nemours & Co., Inc.
Tissue-Tek® is a registered trademark of Sakura-Finetek.
TRI Reagent® is a registered trademark of Molecular Research Center, Inc.
TRIzol® is a registered trademark of Molecular Research Center, Inc.
TurboBlotter™ is a trademark of Whatman Ltd.
Tween® is a registered trademark of ICI Americas, Inc.

Typhoon™ is a trademark of GE Healthcare Life Sciences.

Ultrafree® is a registered trademark of Millipore Corporation.

Vent® is a registered trademark of New England BioLabs.

Whatman® is a registered trademark of Whatman International, Ltd.

Ziploc® is a registered trademark of S.C. Johnson & Son, Inc.

These trademarks and registered trademarks are accurate to the best of my knowledge.
Author

This page intentionally left blank

Glossary

Abundance The prevalence of a particular RNA, or class of RNA molecules, in the cell. The relative amount of a particular messenger RNA (mRNA) species, compared to other samples, is frequently expressed in terms of relative abundance.

Accession Number A numeric identifier of a nucleic acid sequence (genomic DNA, cDNA, or RNA) or an amino acid sequence (protein). The accession number acts as an electronic card catalogue of sorts, facilitating expedient recovery from a database.

Acetylation Introduction of an acetyl group (-COCH₃) to a molecule.

Adduct A molecule to which a chemical group or other molecule has been covalently added, i.e., a chemical addition product.

Adenine A purine that ordinarily base-pairs with thymine in DNA or uracil in RNA.

Adenosine A nucleoside containing the base adenine.

Affinity chromatography A physical separation of molecules based on a biological activity or a biological structure. Examples include base-pairing between the poly(A) tail of some eukaryotic mRNAs with oligo(dT), and the capture of oligohistidine fusion proteins on nickel resin columns.

Alkylation Modification of a molecule by addition of an alkyl group (carbon and hydrogen atoms only, arranged in a linear chain, with the generic formula C_nH_{2n+1}). Alkylation of nucleic acids is often the causative event in the mutation of a gene.

Allele One of two or more alternative forms of a gene.

α particle One of the three forms of natural radiation consisting of two protons and two neutrons.

Alu repeat sequence One of several hundred thousand sequences, about 300 base pairs in length, found dispersed throughout the primate genome. These repetitive sequences are so-named because of the characteristic *Alu* restriction enzyme sites (AGCT) that they contain.

α-Amanitin A bicyclic octapeptide toxin derived from the very poisonous mushroom *Amanita phalloides*. This fungal product differentially inhibits eukaryotic RNA polymerases, and RNA polymerase II is especially sensitive.

Amber codon The translation termination codon UAG.

Amino acids The building blocks of proteins.

AmpliTaq® Recombinant form of the naturally occurring thermostable *Taq* DNA polymerase from the organism *Thermus aquaticus*.

Annealing The base-pairing of complementary polynucleotides to form a double-stranded molecule.

Anion an atom or molecule that has a negative charge. Anions form by the acquisition of electrons by an atom.

Antigen Any molecule or portion thereof that can induce the synthesis of an antibody or otherwise provoke an immune response.

Antiparallel The manner in which two complementary polynucleotides base-pair to

one another; the 5' and 3' ends of each molecule are reversed in relation to each other, so that the 5' end of one strand is aligned with the 3' end of the other strand. Antiparallel base pairing accompanies the formation of double-stranded DNA (DNA:DNA), double-stranded RNA (RNA:RNA), and DNA-RNA hybrids (DNA:RNA).

Antisense RNA (also (-)RNA) Any RNA molecule that is complementary (opposite polarity) to the naturally occurring sequence of mRNA. Antisense RNA is capable of hybridization with mRNA on a filter, *in situ*, or in solution, thereby forming an extremely stable double-stranded RNA.

AOI (area of interest) A contiguous subset of pixels defined within an image, which may be arranged in any polygonal shape, and is used to isolate a subset from the rest of the image.

Aptamer Small DNA or RNA oligonucleotide that can bind a specific ligand (target molecule) by folding into a characteristic three-dimensional shape.

Argonaute A highly conserved family of proteins that are essential components of gene silencing via RNA interference (RNAi). Among the multiple functions of the argonaute proteins is a role in the formation of the RNA-induced silencing complex (RISC) responsible for the mRNA degradation.

Array See microarray and macroarray.

Attenuation Method of regulation of transcription termination in certain prokaryotic operons.

Autoradiograph (also autoradiogram) A photographic record of the spatial distribution of radiation in an object or specimen. It is made by placing the object very close to a photographic film or emulsion.

Autoradiography The process by which the decay of radiolabeled molecules is recorded on X-ray film or other photographic emulsion.

Autosomes Non-sex chromosomes. In humans, chromosome pairs 1–22 are known as the autosomes.

Bacteriophage A virus that infects and is propagated in a bacterial host, often for cloning purposes. Among the best characterized and widely exploited are genetically modified derivatives of the wild type λ bacteriophage.

Base pairing The formation of hydrogen bonds between the nitrogenous bases of two nucleic acid molecules.

β particle An elementary particle emitted from a nucleus during radioactive decay and which has a single electrical charge; it is one of the three forms of natural radiation. A negatively-charged β particle is identical to an electron, and β particle emission is a form of ionizing radiation. A positively-charged β particle is known as positron.

Betaine PCR chemical adjunct that reduces 2° structure in GC-rich regions, thereby improving amplification of DNA. Standard stock solution is 5 M and is used at a working concentration of 0.8–1.6 M.

Bioinformatics The marriage of computers, information technology, and biotechnology. Bioinformatics practitioners are able to sort, archive, and retrieve enormous amounts of nucleic acid and protein sequence information.

Biotechnology (see also genetic engineering) The use of microorganisms, plant-, and animal cells to produce useful materials, such as food, medicine, and other compounds.

Biotin A small water-soluble vitamin, also known as vitamin B7 and vitamin H, which can be used to label nucleic acids for a variety of purposes, including non-isotopic hybridization detection. Alternatively, biotinylated compounds may be isolated by affinity capture using streptavidin.

Bit The smallest unit of information recognized by a computer. A bit, short for binary digit, has a value of 0 or 1. A pixel is represented by one or more computer bits.

The number of bits per pixel determines directly the number of colors or gray shades that can be represented.

Bit depth The number of bits used to represent one pixel value, such as intensity of color. Also referred to as *pixel depth* and *bits-per-pixel (BPP)*.

Bit map A two-dimensional array of image pixels used to represent or store an image. Each cell in the array contains a value that describes a portion of an image in terms of its color.

BLAST (Basic Local Alignment Search Tool) A well-known and widely utilized bioinformatics resource that allows users to search for specific nucleic acid and protein sequences in GenBank® via the internet.

BMP (Bit-map) A digital image file format.

BPP (Bits per pixel) Describes the depth of an image, i.e., the number of bits that describe a pixel in terms of its color. Bi-level images are only 1 BPP while true color images are 24 BPP.

Brightness The amount of white in an image. The brighter the image, the more white it contains. As brightness is increased, each color in the image is shifted more toward white.

Buoyant density A measure of the ability of a substance to float in a standard fluid. For example, differences in the buoyant densities of RNA and DNA allow them to be separated in a gradient of cesium chloride (CsCl) or cesium trifluoroacetate (CsTFA).

CAAT box A conserved regulatory sequence located about 75 base pairs upstream (-75) from the transcriptional start site of eukaryotic genes. The CAAT box is part of the eukaryotic promoter.

Canonical base pairs The usual hydrogen bonding between adenine and thymine, cytosine and guanine, and adenine and uracil.

Cap The structure found at the 5' end of eukaryotic mRNA molecules, including most

eukaryotic viral mRNAs, consisting of an inverted (5'-5') linkage between the first two nucleotides. The 5'-cap is added enzymatically to the 5' end of hnRNA immediately after transcription while the transcript is still in the nucleus. Whereas all caps display the requisite N7-methylguanosine (m⁷G) motif, the variant forms of the 5' cap are Cap 0 (found in unicellular eukaryote), Cap 1 (usually one additional methyl group; most common form), and Cap 2 (usually two additional methyl groups; less commonly observed than Cap 1).

Cation an atom or molecule that has a positive charge. Cations form by the loss of electrons from an atom.

cDNA (complementary DNA) DNA synthesized *in vitro* by reverse transcriptase from an RNA template. cDNA may be single-stranded or double-stranded, as required by the parameters governing a particular assay. cDNA represents a permanent biochemical record of the cellular biochemistry and also provides a means by which that record can be propagated indefinitely.

Central dogma The long held belief that the flow of genetic information in all cells began with the transcription of DNA, the blue print, into an mRNA intermediary, followed by synthesis of the mRNA-encoded protein. The universality of the central dogma was disproven upon discovery of retroviral reverse transcriptase.

Cesium chloride (CsCl) A dense salt used for isopycnic separation of nucleic acids. CsCl is used to partition macromolecules based on differences in buoyant density. The most common applications include purifying plasmids from bacterial cell lysates, and pelleting eukaryotic mRNA with concomitant banding of DNA above the pellet.

Cesium trifluoroacetate (CsTFA) A dense salt used for isopycnic separation of nucleic acids. CsTFA is capable of forming steeper gradients than CsCl. Consequently, RNA can either be banded or pelleted depending on the preference of the investigator.

Chaotropic Biologically disruptive, i.e., causing cellular chaos. Chaotropic lysis buffers disrupt cellular and subcellular membranes and destroy enzymatic activity on contact.

Chargaff's Rule The predication that under ordinary circumstances, the canonical base pairs will form (A::T, A::U, C::G). Knowing that canonical base-pairing is strongly favored over the formation of non-canonical base pairs, one can predict the base composition of double-stranded DNA or RNA if the amount of only one of the bases is known.

Chemiluminescence A non-isotopic hybridization detection technique. Chemiluminescence is the production of visible light by an enzyme-mediated chemical reaction.

Chromatid In metaphase chromosomes, one of two plainly visible distinct subunits. Chromatids are held together by a common centromere until they separate during anaphase.

Chromatin The complex of genomic DNA and protein found in the nucleus of a cell in interphase. At the onset of mitosis, the chromatin condenses into recognizable chromosomes.

Chromosome jumping A technique very similar to chromosome walking with the exception that very large (*Not* I) fragments of DNA (>100 kb) under investigation are purified by pulse-field gel electrophoresis, ligated into a plasmid, and then redigested with another enzyme (*Bam* HI). In so doing, two segments of genomic DNA which were once greatly separated are now within easy sequencing proximity. This type of subcloning constitutes more of a "jump" than a "walk." This method is particular useful in areas of the genome that contain highly repetitive sequences that would be difficult to align by traditional chromosome walking. *See* Poustka, A. *et al.* (1987). Construction and use of human chromosome jumping libraries from *NotI*-digested DNA. *Nature* 325, 353.

Chromosome walking The systematic isolation of a set of clones containing overlapping

DNA fragments which collectively constitute a specific genomic region. The process is initiated by recovery of a unique clone using a sequence-specific probe. When the clone is retrieved from a library, the process of chromosome walking proceeds by subcloning and rehybridization to the library: in each successive hybridization the next probe used corresponds to the 3' end or 5' end of the clone previously recovered from the library. The resulting series of overlapping clones permits locus characterization within, upstream, and downstream of a particular locus. Thus, this approach allows the investigator to "walk" along a chromosome in both directions. This is the method by which the gene associated with Duchenne muscular dystrophy was originally identified. Shotgun DNA sequencing can be used as an alternative to this method when the whole genome sequencing is involved. *See* Monaco, A.P. *et al.* (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* 323, 646-650.

Cistron A discrete region or specific sequence of DNA. The term cistron is an older name for what is commonly called a gene.

Clone (noun) A collection of genetically equivalent cells or molecules.

Clone (verb) A series of manipulations designed to isolate and propagate a specific nucleic acid sequence, or cell, for characterization, storage, or further amplification.

Clone bank Older terminology for what are currently known as cDNA libraries and genomic DNA libraries.

Coding strand In double-stranded DNA, the strand that has the same sequence as the resulting RNA (except for the substitution of uracil for thymine). The coding strand alone is reported when publishing a gene sequence and is always written 5'→3', from left to right.

Codon A group of three nucleotides found in the coding portion of mRNA. The precise order of nucleotides specifies the identity and

placement of an amino acid into a nascent peptide. Punctuation (start and stop) is also codon-specified.

Colinearity The manner in which the linear order of nucleotides, read 5'→3', dictates the order of amino acids in the resulting protein, read from the amino terminus to the carboxy terminus. The order of codons specifies the order of amino acids.

Competitive PCR A very sensitive method for quantification of transcript abundance by the inclusion of an additional template DNA sequence in the same reaction tube as an experimental sample, thereby competing for primers, dNTPS, and the "attention" of the polymerase. By varying the amount of the competitor DNA (occasionally known as a DNA mimic), a dilution will be identified in which the concentration of target and mimic are equivalent, resulting in equimolar amplification and precision quantification.

Compression A mathematical technique that allows an image to be stored with less memory. Redundancies in the internal representation of the image are identified and given a code which then replaces the data redundancies.

Consensus sequence A composite nucleotide- or amino acid sequence that has been deduced from the examination and alignment of multiple molecules that share this sequence similarity. Succinctly, a consensus sequence is essentially the "average" of several similar sequences. In cases where individual nucleotides or amino acids differ from the majority of the sequences being considered it is customary to indicate the ambiguity by either assigning a percentage (A_{90} , meaning that adenosine occurs 90% of the time in a particular position) or by listing the observed permutations ($ATCG^C_A TGTAT$).

Conserved sequence A DNA sequence which exhibits little or no variation from one organism to the next.

Constitutive gene expression Continuous interaction of RNA polymerase with the

promoters of specific genes not subjected to additional levels of regulation. Such genes are often expressed continually at low or basal levels, and are sometimes referred to as genes with housekeeping functions.

Contig A contiguous stretch of DNA that has been assembled *in silico* from a series of smaller overlapping cloned genomic or cDNA segments.

Contrast The sharpness of an image. The higher the contrast in an image, the larger is the difference between the darkest and lightest areas in an image.

Cosmid A standard molecular biology vector in which the λ bacteriophage *cos* sequences are joined to a plasmid. Thus, after ligation of an insert (foreign) DNA, the cosmid is packaged into λ bacteriophage for infection into a bacterial host, after which the cosmid begins acting like a plasmid. Cosmids are able to accept much larger inserts than traditional plasmids.

Countertranscript RNA (ctRNA) Another name for antisense RNA that is encoded in the same genomic region, but in the opposite orientation, as its functional sense RNA counterpart. Consequently, sense RNA and ctRNA are able to form areas of dsRNA and, in so doing, suppress transcription initiation elsewhere. ctRNA is often discussed in the context of prokaryotic gene expression.

cRNA (complementary RNA) Usually another name for antisense RNA, i.e., a transcript which is complementary to mRNA. Occasionally, however, cRNA can also mean a transcript with a similar sequence to mRNA and which can act as a competitor in an amplification reaction.

C_T (Threshold cycle) The cycle in real-time PCR when accumulating product crosses a predetermined threshold level and becomes visible to the system. C_T is the basic unit of measurement in real-time PCR.

Curie (Ci) A basic unit for measuring radioactivity in a sample. One curie is equivalent

to 3.7×10^{10} Becquerel (i.e., disintegrations per second).

Cy3 (indocarbocyanine) A fluorescent water-soluble dye (absorption 550 nm; emission 570 nm) often used in combination with Cy5 to label DNA for various applications, especially microarray-based assays. Both Cy3 and Cy5 are more stable, brighter fluorophores than other traditional fluorescent compounds and produce less background.

Cy5 (indodicarbocyanine) A fluorescent water-soluble dye (absorption 650 nm; emission 670 nm) often used in combination with Cy3 to label DNA for various applications, especially microarray-based assays. Both Cy5 and Cy3 are more stable, brighter fluorophores than other traditional fluorescent compounds and produce less background.

Cycloheximide A drug that inhibits protein synthesis in eukaryotic cells. No effect on prokaryotic cells.

Cytidine A nucleoside containing the base cytosine.

Cytogenetics That branch of genetics focusing on the structure and behavior of chromosomes, especially during mitosis and meiosis.

Cytoplasm The cellular contents found between the plasma membrane and the nuclear membrane.

Cytosine A pyrimidine that usually base-pairs with guanine in DNA.

Degenerate codon One of two or more codons that specifies the same amino acid.

Denaturation (of nucleic acids) Conversion of DNA or RNA from a double-stranded form to a single-stranded form. This can mean dissociation of a double-stranded molecule into its two constituent single-strands, or the elimination of intramolecular base-pairing that is commonly associated with RNA.

DNA (deoxyribonucleic acid) A polymer of deoxyribonucleoside monophosphates, assembled by a DNA polymerase. *In vivo*, DNA is

produced by the process known as replication. DNA can also be synthesized using a variety of *in vitro* methods, such as PCR.

Dicer A dsRNA-specific endonuclease that produces short interfering RNA (siRNA) and microRNA (miRNA) cleavage products from longer templates. These shorter molecules are intimately associated with the RNA interference (RNAi) pathway in the cell.

Dideoxynucleotide A modified DNA nucleotide that exhibits a 3'-H, rather than the usual 3'-OH. Incorporation of a dideoxynucleotide prevents the further elongation of nascent strands of DNA during sequencing reactions due to their lack of the 3'-OH needed for phosphodiester bond formation with what would be the next nucleotide in the chain.

Diethyl pyrocarbonate (DEPC) A chemical used to purge reagents of nuclease activity. DEPC is carcinogenic and should be handled with extreme care. Nuclease-free H₂O and other reagents are also readily available from most vendors, thereby precluding the need to use DEPC for this purpose in the laboratory.

Differential Display PCR (DDPCR) *also* mRNA differential display. A method for identification of uniquely transcribed sequences among two or more RNA populations. DDPCR is a PCR-based method that utilizes large combinations of relatively short primers to ensure amplification of all cDNA species within a sample. Electrophoretic comparison of the PCR products of each reaction shows products of identical molecular weight when a transcript is common to the biological samples under investigation; a band in only one lane is observed if gene expression has been induced or repressed. Detection of up- and downregulated genes is also possible using this method.

Differentiation The process of biochemical and structural changes by which cells become highly specialized in form and function.

DMSO (dimethyl sulfoxide) A reagent used in conjunction with glyoxal to denature RNA prior to electrophoresis. DMSO can also be

used to lower the T_m (melting temperature) of DNA duplexes to support long-range PCR.

Diploid Having two complete sets of chromosomes (two of each chromosome). Compare with *haploid* (one of each chromosome) and *triploid* (three of each chromosome).

Directional cloning Unidirectional insertion of a DNA molecule into a vector accomplished by placement of different sequences or restriction enzyme sites at the ends of double-stranded cDNA or genomic DNA molecules.

DNA polymerase An enzyme capable of synthesizing DNA from a DNA template. There are five known DNA polymerases in *E. coli* and nine different DNA polymerases are known in eukaryotic cells. Each polymerase has a specific role in DNA replication or repair, and exhibits a characteristic combination of distinct activities, namely: 5'→3' polymerase, 5'→3' exonuclease, and 3'→5' exonuclease.

Dot blot A membrane-based technique for the quantification of specific RNA or DNA sequences in a sample. The sample is applied directly onto a filter by vacuum filtration through a manifold in either a dot or slot configuration (*see also* slot blot). Dot blot analyses lack the qualitative component associated with electrophoresis-based techniques and are generally considered low sensitivity assays.

Dounce homogenizer A device for tissue disruption that consists of a glass tube with an accompanying tightly-fitting pestle. This is a hand-operated device that is generally less disruptive than electronic homogenizers.

Downstream Sequences in the 3' direction (further along in the direction of expression) from some reference point. For example, the initiation codon is located *downstream* from the 5' cap in eukaryotic mRNA.

Duplex A double-stranded molecule or portion thereof that forms by the base-pairing of two complementary polynucleotides.

Eberwine linear amplification A widely used process for the linear amplification of

antisense RNA (cRNA) from cDNA templates, most often for microarray-based applications. This technique minimizes the distortion of abundance relationships among transcripts in the original sample because it lacks the amplification bias inherent in PCR.

Electropherogram A tracing or plot of the results of a DNA sequencing reaction.

Electrophoresis A type of chromatography in which macromolecules such as proteins and nucleic acids are resolved through a matrix based on their size and net charge.

Electrophoretogram A photograph of a gel made after electrophoresis, which records the spatial distribution of macromolecules within the gel. An electrophoretogram is widely referred to as "picture of the gel".

ELISA (enzyme-linked immunosorbent assay) An assay that detects an antigen-antibody complex via an enzyme reaction, thereby producing a visible result such as a color change (chromogenic detection) or measurable fluorescence.

Enhancer element A short regulatory DNA sequence that can increase the activity of some eukaryotic promoters, resulting in elevated levels of transcription. Enhancer elements exert their influence relatively independently of orientation and proximity to target promoters.

Enrichment Any manipulation of RNA that results in an increase in the statistical representation of one or more RNA subpopulations as a percentage of all RNAs synthesized by the cell. For example, selection of poly(A)⁺ RNA results in an enrichment of mRNA. An enrichment strategy may also involve the physical manipulation of cells before lysis to superinduce one or more species of RNA: serum stimulation of quiescent cells, for example, results in an induction (enrichment) of proliferation-specific message as cells reenter the cell cycle.

EST (expressed sequence tag) Partial sequences from selected cDNA clones which

match at least a portion of the coding region of a gene. Public databases contain immense amounts of sequencing data useful for the identification of ESTs.

Ethidium bromide (EtBr) A planar, intercalating agent used to visualize nucleic acids, both DNA and RNA. This dye emits a bright orange fluorescence when UV irradiated; thus, gels that contain nucleic acids can be photographed for future reference. The standard ethidium bromide stock solution is 10 mg/ml in water and the standard staining concentration is 0.5 to 1.0 μ g/ml.

Exon A portion of a eukaryotic gene represented in the mature mRNA molecule. Exons may or may not be translated. Exons are separated by introns prior to the splicing of precursor hnRNA molecules in the nucleus.

454 sequencing Novel sequencing platform described as massively parallel pyrosequencing. This technology enables the user to sequence DNA, without the need for cloning, at a rate of millions of bases per hour.

File format The specific organization method by which an image is stored on a disk, based on image class, compression type, and half-tone pattern. TIFF, BMP, JPEG, and GIF are examples of file formats.

Formaldehyde A commonly used denaturant of RNA and inhibitor of RNase. It should be handled with care in a chemical fume hood. Formaldehyde is a liver carcinogen and, as a known teratogen, should be avoided by expectant mothers.

Formamide A commonly used organic solvent/denaturant used to lower the melting temperature (T_m) of double-stranded duplexes; for this reason it is often included in hybridization reactions to maintain stringency at a temperature lower than might otherwise be required. It is often used with formaldehyde to denature RNA before electrophoresis. Formamide is a carcinogen and, as a known teratogen, should be avoided by expectant mothers.

Fosmid An f-factor cosmid. Cosmids and fosmids are similar in structure to plasmids, but have the ability to carry significantly larger inserts, often on the order of up to 45–50 kb of DNA. The upper limit for many plasmid-borne inserts is about 10 kb DNA. Because of their great size a bacterial cell is able to acquire and replicate only one fosmid, making fosmids less desirable cloning vehicles compare to plasmids. The use of fosmids has been largely supplanted through the development and implementation of bacterial artificial chromosomes (BACs) for cloning purposes.

Frame grabber Hardware unit, essentially a card inserted into one of the slots in the back of a computer that permits the use of a video camera to import an image directly into image analysis software.

Functional genomics Study of the function of genes by examination of the levels of RNA (and protein) in the cell and how these levels are modulated as the cell responds to change. Succinctly, it is the study of how the genome responds. *See also* transcriptomics.

Gamma A non-linear logarithmic contrast correction factor used to adjust the contrast in dark areas of an image.

γ radiation The most energetic of the three forms of natural radiation. Unlike α and β subatomic particles, γ rays consists solely of electromagnetic radiation.

Gene The unit of heredity. A gene is a discrete sequence of chromosomal DNA that ultimately encodes the instructions for the synthesis of a polypeptide.

Genetic engineering *Also* recombinant DNA technology, biotechnology, molecular biology. The technology by which genes, or pieces thereof, are moved from one organism to another. Genetic engineering has grown into an industry with far-reaching implications in virtually all areas of science and society.

Genome The entirety of chromosomal DNA found in a cell. In some applications, it may be useful to distinguish nuclear genomic DNA from the mitochondrial genome.

Genomic DNA Chromosomal DNA. The mitochondrial genome is designated as mtDNA, the chloroplast genome is designated ctDNA, and nuclear DNA is generically referred to as genomic DNA.

Genomics Study of the structure and organization of genes. Genomics is intimately linked to data derived from DNA sequencing.

Genotype The actual collection of genes and alleles present in a cell or organism.

GIF (graphics interchange format) A bit-mapped color graphics format that supports 256 colors. Digital images of company logos, line art, and related applications are frequently stored as GIF files.

Glyoxal (OHCCHO) A reagent that has been used to denature RNA prior to electrophoresis. It is often used in conjunction with DMSO and is less popular than the formaldehyde denaturing system.

Gray level The brightness value assigned to a pixel in gray scale images. In an 8-bit image, this value ranges from 0 to 255, specifically from black, through shades of gray, to white.

Guanine A purine that usually base-pairs with cytosine.

Guanosine A nucleoside containing the base guanine.

Haploid Having one complete set of chromosomes (one of each chromosome, as found in gametes). See also *diploid* and *triploid*.

Hapten A small molecule, not antigenic by itself, that can act as an antigen when conjugated to a larger antigenic molecule, usually a protein.

Heat Map A graphical two-dimensional data display in which data values are often represented in color. This is a common method for summarizing the up- and downregulation of a number of genes from a number of samples.

hnRNA (heterogeneous nuclear RNA) Precursor mRNA; the primary product of eukaryotic transcription by RNA polymerase II.

Histone proteins DNA binding proteins that facilitate supercoiling of chromatin into recognizable chromosomes. Histone proteins are among the most highly conserved among all eukaryotic genes and are the core of nucleosome formation. Histone mRNAs are often assayed as housekeeping genes.

Hogness box (also TATA box) Older term for the conserved TATAA-motif located about 25 bp upstream (-25) from the transcriptional start site of eukaryotic genes. The Hogness box is a *cis*-acting regulatory component of many eukaryotic promoters for RNA polymerase II.

Holoenzyme A complete enzyme (as opposed to a cleavage product).

Hot start PCR A chemical or procedural modification that prevents synthesis of any PCR product until after an initial period of high temperature denaturation. Hot start PCR is frequently used to maximize the sensitivity of PCR by eliminating the synthesis of non-specific product.

Housekeeping gene A gene that is expressed, at least theoretically, at constant levels in all cells, the products of which are required to maintain normal cell function or viability. Because of their purported invariance, assay of transcription of these sequences is often performed to demonstrate that an *overall* change in gene expression has *not* occurred, in the context of an experimental manipulation.

Hybridization The formation of hydrogen bonds between two nucleic acid molecules that demonstrate some degree of complementarity. The specificity of hybridization is a direct function of the stringency of the system in which the hybridization is being conducted.

Hydrogen bonding The highly directional attraction of an electropositive hydrogen atom to an electronegative atom such as oxygen or nitrogen. This is the manner of interaction of complementary bases during nucleic acid hybridization. Hydrogen bonds

are relatively weak and are easily broken by increasing stringency. *See also* base pairing.

Hydroxylation Chemical reaction involving the addition of one or more -OH groups, most often by enzymes known as hydroxylases. Protein hydroxylation is usually directed at proline residues and is one several posttranslational modifications that can influence protein form and function.

Hyperpolymer A collection of labeled probe molecules that have hybridized to a target sequence and with each other, in an overlapping fashion, such that a tail of labeled molecules is attached to the target sequence. This phenomenon occurs when breakage of a nucleic acid backbone occurs during labeling, as in nick translation, or as a result of failure to seal the backbone during probe synthesis, as in random priming. The effect is an amplification of the signal, which translates into shorter autoradiographic exposure time, although a slight loss in resolution is observed.

Image A two-dimensional array of data elements, each of which is related to a color or an intensity. This information can be stored in any of several file formats and can be evaluated using image analysis software.

Image analysis An electronic method for the digital capture and storage of an image, accompanied by automated measurement of parameters such as molecular weight, mass, relative abundance, and optical density of various objects in the image (e.g., bands on a gel).

Inosine (I) An unusual nucleoside containing the base hypoxanthine. Inosine is sometimes used in the synthesis of degenerate or "guess-mer" oligonucleotide primers because it is capable of base-pairing with adenine, cytosine, and thymine.

In silico Through the use of a computer. For example, using bioinformatics software to predict the location of a transcription start site.

In situ hybridization A method for localizing gene expression in tissue. RNA or DNA

probes are used for hybridization to mRNA present within a very thin tissue section. By maintaining natural histological geometry, changes in gene expression can often be mapped to a few cells within a tissue.

Intron An intervening DNA sequence that interrupts the coding sequences (exons) of a gene. Introns are transcribed and represented in hnRNA, though they are spliced out, in an apparently systematic fashion, during the RNA maturation.

Ionizing radiation Any radiation that displaces electrons from atoms or molecules, resulting in the formation of ions. α , β , and γ radiation are all forms of ionizing radiation.

Isoprenylation Linkage of lipid groups to cysteine residues. This is one of several possible posttranslational modifications associated with protein maturation.

Isotope One of two or more atoms with the same atomic number but different atomic weights, i.e., different atomic forms of the same element.

JPEG (joint photographic experts group) A widely used electronic compression method used to digitize color images and photographs.

Klenow enzyme The large fragment of *E. coli* DNA polymerase I, generated by cleavage of the holoenzyme with the protease subtilisin or obtained by cloning. The Klenow fragment retains the 5'→3' polymerase and 3'→5' exonuclease activities but lacks the often troublesome 5'→3' exonuclease activity associated with the intact enzyme. *See also* DNA polymerase.

Leader sequence The non-translated portion of mature mRNA located between the 5' cap and the initiation codon.

Library A collection of clones that partially or completely represent the complexity of gene sequences from a specific biological source. A library consists of cDNA or genomic DNA molecules that have been ligated to a suitable vector and which can

be propagated indefinitely. Archived DNA library members may be selected or retrieved from the library by nucleic acid hybridization or, in the case of expression vectors, by antibody recognition.

Ligase Any enzyme that joins together nucleic acid molecules. Ligases may be thought of as “molecular paste” and, in the case of DNA ligase, reverse the action of restriction endonucleases (molecular scissors). T4 DNA ligase is used to join double-stranded DNA molecules and T4 RNA ligase is used to join single-stranded RNA or single-stranded DNA molecules.

Ligase Chain Reaction (LCR) A DNA amplification technique that was developed primarily to identify point mutations. The most common form of LCR involves the ligation of two primer molecules which flank a putative mutation site, resulting in a single, longer primer which is needed to support subsequent amplification. If a point mutation has occurred, then the two shorter oligonucleotides will not be ligated because of a mismatch involving the end of one of the two primers and amplification, therefore, will not be supported.

Linear RNA Amplification *See* Eberwine linear amplification.

Locked nucleic acid (LNA) A modified nucleotide with an extra bridge joining the 2' oxygen and 4' carbon atoms. The inclusion of LNA in oligonucleotide primers or probes significantly increases the T_m , thereby reducing the incidence of non-specific product formation while increasing the discriminatory power of the primers.

Locus (plural is loci) The precise location of a particular gene, and any possible alleles, on a chromosome.

Long-Range PCR Modified PCR amplification procedure by which very large PCR products are generated. Reaction modifications usually include a quasi-alkaline reaction buffer, higher nucleotide concentrations,

a pair of thermostable enzymes, higher T_m primers, and a modified cycling profile.

Macroarray A nylon filter onto which multiple gene sequences have been immobilized. Macroarrays are able to hold only a small fraction of the number sequences that can be printed on a microarray. Macroarrays are often themed (cancer, apoptosis, mouse, human, and so forth) and lack the sensitivity associated with microarray technology.

MALDI-TOF (matrix-assisted laser desorption ionization-time of flight) A form of mass spectrometry used to sequence peptide fragments for proteomics applications.

Marker A generic term that can refer to any allele of interest in an experiment. In addition, “marker” can refer to a molecular weight standard.

Meiosis Cell division responsible for the production of haploid gametes (sex cells).

Melting temperature *See* T_m .

Mendel (Father Gregor Mendel) A Catholic priest who, in 1865, first postulated the notion of the gene. Mendel is universally recognized as the father of classical genetics, also known as Mendelian genetics in his honor.

mRNA (messenger RNA) The mature product of RNA polymerase II transcription. In eukaryotic cells, mRNA is derived from hnRNA. In conjunction with the protein translation apparatus, mRNA is capable of directing the synthesis of the encoded polypeptide. In prokaryotic cells mRNA may carry instructions for the synthesis of two or more proteins.

Metabolome All of the metabolites (small, non-polymerized compounds and monomers) that are present in a biological sample and are able to participate in general metabolic reactions. The composition of the metabolome at any particular moment is a chemical fingerprint of sorts and an indicator of the physiology of the cell.

Metabolomics The science of determining the variety, number, and abundance of low molecular weight compounds (metabolites) in a sample or in a cell or tissue sample. It is widely believed that analysis of the metabolome may be indicative of the role of genes with as-of-yet unknown function(s). Some of the methods used in metabolomics (and metabonomics, below) include nuclear magnetic resonance (NMR), gas chromatography (GC), liquid chromatography (LC), two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), and mass spectrometry.

Metabonomics The science of metabolomics in humans and other higher animals only. More specifically, metabonomics attempts to characterize and classify the metabolic response of cells to drugs, the onset of specific diseases, and changes in the local environment.

Methylation The addition of a $-CH_3$ (methyl) group to a variety of substrates.

Methylmercuric hydroxide (also methyl mercury) An extremely toxic reagent no longer used to denature RNA prior to electrophoresis or for other chromatographic applications. Methylmercuric hydroxide should always be avoided in favor of other less dangerous denaturation options.

Microarray A glass chip printed with tens of thousands of different DNA sequences for high-throughput global analysis of gene expression.

miRNA (micro RNA) Very short abundant transcripts that play a role in the regulating gene expression by either transiently suppressing translation of an mRNA molecule or by directing its cleavage.

Mismatch A position in a double-stranded molecule where the nucleotides do not base-pair. In order for mismatches to be tolerated, the temperature of annealing must be sufficiently below the melting temperature (T_m); at the T_m only perfectly matched duplexes are stable. The location and context of mismatching have profound ramifications with respect to primer annealing, and the identity

of the resulting product(s), when performing PCR.

Mitosis Cell division in somatic cells. Mitosis is responsible for producing genetically identical diploid copies ($2n \rightarrow 2n$) from a parental cell.

Molecular beacons Fluorescent hybridization probes that are used in certain real-time PCR applications.

Molecular biology The discipline dealing with the study of gene organization and gene expression at the RNA and DNA levels. Molecular biology techniques involve making temporary or permanent changes to the structure of nucleic acid molecules and isolating specific sequences for further analysis. Somewhat synonymous with biotechnology, genetic engineering, and recombinant DNA methodology.

Monocistronic Term used to describe mRNA molecules that encode only one polypeptide. Monocistronic mRNAs are commonly associated with eukaryotic cells.

MOPS 3-[N-morpholino]propanesulfonic acid; a key component of the buffering system used in conjunction with formaldehyde-agarose gel electrophoresis of RNA. $10\times$ MOPS Buffer = 200 mM MOPS, pH 7.0; 50 mM sodium acetate; 10 mM Na_2 -EDTA, pH 8.0.

MSDS (Material Safety Data Sheet) A document that describes many of the physical and chemical properties of a substance, along with instructions for proper handling, storage, containment, and disposal. An MSDS also contains toxicity data and first aid information, in the event of exposure to or ingestion of the substance. MSDS information should be maintained on file in the laboratory, with ready access at all times.

Multiplex PCR Use of multiple primer pairs for the amplification of one or more targets in the same PCR reaction tube.

NASBA (nucleic acid sequence-based amplification) An isothermic process of

“self-sustained sequence replication” used primarily for molecular diagnostics.

Nascent protein A protein molecule which has not yet been completely synthesized by the process of translation.

Nascent RNA An RNA molecule which has not yet been completely synthesized by the process of transcription.

Normalization Any of a number of measurements or manipulations made to allow meaningful comparison among samples in a series or on a filter membrane. Normalization typically involves ensuring equal amounts of template material in an assay, or standardizing data against a baseline signal.

Northern analysis (also Northern blotting) A technique for transferring electrophoretically chromatographed RNA from an agarose gel matrix onto a filter paper, for subsequent immobilization and hybridization. The information gained from Northern analysis is used to qualitatively and quantitatively assess the expression of specific genes. The Northern analysis is characteristically lacking in sensitivity by comparison with other transcription assays.

Nuclear runoff assay A method for labeling nascent RNA molecules in the nucleus as they are being transcribed. In some circles, this technique is also known as the nuclear runoff assay. The rate at which specific RNAs are being transcribed can then be assayed based on the degree of label incorporation. *Compare with steady-state RNA.*

Nuclease protection assay A method for mapping and/or quantifying the abundance of transcripts. In general, hybridization between probe and target RNA takes place in solution, followed by nuclease digestion of all molecules or parts thereof that do not actually participate in duplex formation. Nucleic acid molecules locked up in a double-stranded configuration are relatively safe, i.e., protected, from nuclease degradation. The undigested RNA:RNA or RNA:DNA hybrids are then electrophoresed or precipitated for quantitation.

Nucleoside A molecule consisting of a five-carbon sugar (ribose or deoxyribose; pentose) linked to a nitrogenous base (adenine, cytosine, guanine, thymine, or uracil).

Nucleotide A molecule consisting of a five-carbon sugar (ribose or deoxyribose; pentose), a nitrogenous base (adenine, cytosine, guanine, thymine, or uracil), and at least one phosphate group. Nucleotides are the building blocks used to assemble both DNA and RNA.

Ochre codon The translation termination codon UAA.

Oligonucleotide A short, artificially synthesized, single-stranded DNA molecule that can function as a nucleic acid probe or a PCR primer. *Oligonucleotide* can also refer to a short fragment of RNA.

Opal codon The translation termination codon UGA.

Ortholog A gene in one organism that has a similar function to a gene found in another organism. Identifying orthologs, known more formally as orthologous genes, implies some degree of evolutionary relatedness.

Palindromic sequence A short, double-stranded DNA sequence in which the order of bases is identical when read 5' to 3' on each strand.

PCI Standard laboratory symbol for a mixture of phenol:chloroform:isoamyl alcohol (25:24:1).

PCR (polymerase chain reaction) A primer-mediated enzymatic process for the systematic exponential amplification of minute quantities of specific genomic or cDNA sequences. This technique has revolutionized molecular biology in the past 20 years. It has the advantage of being a very sensitive technique that can be performed very quickly in the lab. Using this technique, an investigator can isolate and amplify one sequence from an extremely heterogeneous pool, produce novel templates for gene expression studies, prepare DNA fingerprints, and study

DNA-protein interactions to name but a few applications.

Pentose A five-carbon sugar, either ribose or deoxyribose. A pentose sugar is one of the components of a nucleotide and a nucleoside.

Pharmacogenetics That discipline concerned with the analysis of gene regulation in response to drugs.

Pharmacogenomics That discipline concerned with the interactions of every part of the genome and, therefore, the response of a patient to drugs. The use of this term is more inclusive than the term *pharmacogenetics*, which is more concerned with the response of individual genes. The pharmacogenomics discipline is concerned with cellular pathway interactions with the goal of creating personalized drugs.

Pharmacokinetics The study of absorption, processing, and clearance of drugs from the body over a particular time line.

Phosphate-buffered saline (PBS) An isotonic salt solution frequently used to wash residual growth medium from the cell monolayer. One liter of calcium-magnesium-free PBS is prepared by mixing 8 g NaCl, 0.2 g KCl, 1.15 g Na₂HPO₄, 0.2 g KH₂PO₄. PBS should then be autoclaved and aseptically handled in order to maintain tissue culture sterility.

Phosphorimaging A filmless process by which radioactive decay events, fluorescence emission, and chemiluminescence may be detected and electronically analyzed. Phosphorimagers provide a much greater linear response than X-ray film.

Phosphorylation The addition of a phosphate group to a target molecule, often by the action of a kinase. Phosphorylation is a widely used mechanism for controlling the function of enzymes and proteins.

Photodocumentation A method for preserving the image of a gel immediately after electrophoresis or after hybridization with a labeled probe. Virtually all gel images are now

captured and stored electronically though older media, including Polaroid film, X-ray film, and thermal paper, persist. *See* image analysis.

Photon A particle that has no mass but is able to transfer electromagnetic energy.

Pixel Picture element, i.e., a single point in a graphic image. Pixels are the resultant smallest units of image digitization.

Plasmid A covalently closed, double-stranded DNA molecule capable of autonomous replication in a prokaryotic host cell. Eukaryotic plasmids have also been developed and are in widespread use. Plasmids can accept DNA inserts from any other biological source, though the insert size is usually limited to about 10 kb. Plasmids generally contain a variety of selectable markers and ancillary sequences for characterization, and sometimes expression, of the insert DNA.

Poly(A)⁺ tail A tract of adenosine residues enzymatically added to the 3' terminus of mRNA by the nuclear enzyme poly(A) polymerase. The addition of the poly(A) tail involves cleavage at the 3' terminus of the primary transcript followed by polyadenylation. The length of the poly(A) tail correlates with the stability of that particular RNA in the cytoplasm, and the poly(A) tail shortens as an mRNA ages.

Polyadenylation signal A highly conserved six-base motif (AAUAAA) instrumental in the efficiency of polyadenylation. The splicing events that precede the actual polymerization of adenosine residues to the 3' end of eukaryotic mRNA molecules are controlled, in part, by the polyadenylation signal. AAUAAA is functionally linked to downstream GU-rich and U-rich regions on the native transcript.

Polycistronic Term used to describe mRNA molecules that contain coding regions representing more than one polypeptide.

Polymorphism A genetic trait that, within a species, is subject to variation. Examples of human polymorphisms include eye color,

blood type, cutaneous fingerprints, and DNA fingerprints.

Polysome An mRNA molecule that is being translated simultaneously by several ribosomes.

Polytron An electronic tissue disruption instrument. Polytron homogenization much more harsh than using a Dounce homogenizer.

Posttranscriptional regulation Any event occurring after transcription termination and that influences any of the subsequent steps involved in the ultimate expression of that gene. Reference to posttranscriptional regulatory events include, but are not limited to, intron removal, polyadenylation, nucleocytoplasmic transport, stability in the cytoplasm, and engagement by the protein translation apparatus.

Posttranslational regulation Any event occurring after synthesis of the primary peptide that influences the subsequent maturation involved in the ultimate expression of the gene. Reference to posttranslational regulation usually refers to the efficiency of the events that modify a peptide including, but not limited to glycosylation, methylation, and hydroxylation.

Precursor RNA (also hnRNA or pre-mRNA) An unspliced RNA molecule; the primary product of transcription produced by eukaryotic RNA polymerase II.

Pribnow box The consensus sequence TATAATG that is part of the transcriptional promoter for prokaryotic genes. It is centered about 10 base pairs upstream (-10) from the transcription start site of bacterial genes. The Pribnow box itself is especially important in the binding of RNA polymerase.

Primer A short nucleic acid molecule that, upon base pairing with a complementary sequence, provides a free 3'-OH for PCR or any of a variety of other primer extension-dependent reactions.

Probe Usually, labeled nucleic acid molecules, either DNA or RNA, used to hybridize

to complementary sequences in a library, or which are among the complexity of different target sequences present in a nucleic acid sample, as in the Northern analysis, Southern analysis, or nuclease protection analysis. Probe specificity can be controlled by raising or lowering stringency.

Processivity The ability of an enzyme such as DNA polymerase to remain bound to its substrate (template) and efficiently produce a long product (complementary strand) before disengaging. For example, *Taq* DNA polymerase is said to be a highly processive enzyme because it travels great distances along a template before dropping off, though this particular thermostable enzyme is well-known for its high error rate.

Promoter A DNA sequence associated with a particular locus to which RNA polymerase and other initiation-associated proteins bind at the onset of transcription. Promoters typically consist of several regulatory elements involved in initiation, regulation, and efficiency of transcription.

Proteome The full complement of proteins produced by a cell at a particular time. Proteome maps are typically generated and assessed by 2-D electrophoresis and other techniques designed to identify, quantify, and characterize the products of translation.

Proteomics The study of expressed proteins; specifically, the study of the protein composition of a cell under certain biochemical conditions, and how the assortment of proteins changes in response to natural and artificial stimuli. This approach facilitates the functional characterization of many proteins simultaneously. Proteomics constitutes a systematic means for correlating protein function and abundance with genome function in normal cells as well as the disease state.

Purine The chemical parent compound of the bases adenine and guanine.

Pyrimidine The chemical parent compound of cytosine, guanine, and uracil.

Quinone Oxidation product of phenol. Quinones compromise the quality of RNA preparations by crosslinking nucleic acid molecules and breaking phosphodiester bonds. Avoid quinones by always preparing and working with fresh phenol. Quinones characteristically impart a pink tint to phenol solutions, indicating that it is time to dispose that particular aliquot.

q.s. Abbreviation for the Latin words *quantum satis*, meaning “sufficient quantity”. The inclusion of q.s. in a protocol or formulation means that the investigator will add enough water to achieve the specified final volume in a reaction tube.

RACE (rapid amplification of cDNA ends) A PCR method for cloning the 5' end (5' RACE) and 3' end (3' RACE) of transcripts. RACE can be performed using a number of formats, the best of which is RNA-ligase mediated RACE (RLM-RACE) for mapping transcription start sites.

Radiochemical A chemical containing one or more radioactive atoms.

Radiolysis The physical breakage of DNA or RNA probe that occurs when these molecules are radiolabeled to extremely high specific activity. Hence, the radioactive decay of the labeled nucleotides is responsible for damaging the probe itself.

Radionuclide An unstable isotope of an element that decays spontaneously and, in so doing, emits radiation.

Real-time PCR A state of the art method for measuring PCR product accumulation as it is produced in each cycle, rather than measuring final product mass at the end of the reaction. Real-time PCR is widely regarded as the premier quantitative PCR-based assay.

Recombinant DNA technology *see* Molecular Biology.

Renaturation (also reannealing) The reassociation of complementary strands of DNA or RNA following high stringency denaturation.

Restriction endonuclease A DNA modification enzyme that cuts (restricts) double-stranded DNA at predictable locations. Thought of as the “molecular scissors”, restriction endonucleases most often cut substrate DNA at palindromic 4-, 6-, or 8-base sequences.

Retrovirus A virus with an RNA genome that propagates via conversion (reverse transcription) of its genetic material into double-stranded DNA.

Reverse Northern Blot A method in which DNA probes immobilized on a filter, often in a dot blot configuration, are hybridized to labeled mRNA or cDNA. “Reverse Northern” are commonly used to confirm differential gene expression data produced by other assays, for example RT-PCR.

Reverse transcriptase (RT) Also known as RNA-dependent DNA polymerase. A retroviral enzyme that polymerizes a complementary DNA (cDNA) molecule from an RNA template. Many types of reverse transcriptase are available, including the very traditional AMV (source: avian myeloblastosis virus) and MMLV (source: Moloney-murine leukemia virus) varieties. Currently, there are available several forms of reverse transcriptase that lack a natural background RNase H activity, thereby making them desirable for cDNA synthesis reactions. Some of the thermostable polymerases that support PCR, such as that from the thermophilic eubacterium *Thermus thermophilus*, have been found to possess a Mn²⁺-dependent reverse transcriptase activity at elevated temperature; these have been used successfully in the one-tube RT-PCR format.

Ribonuclease (RNase) A family of resilient enzymes that rapidly degrade RNA molecules. Control of RNase activity is a key consideration in all manipulations involving RNA, directly or indirectly.

Ribonuclease A (RNase A) A single-strand-specific endoribonuclease with activity directed toward the 3' side of pyrimidine

nucleotides (Py/pN), rendering a 3' pyrimidine phosphate.

Ribonuclease H (RNase H) An enzyme with activity directed against the RNA component of an RNA:DNA hybrid.

Ribonuclease T1 (RNase T1) A single-strand-specific endoribonuclease that cleaves the phosphodiester bond at the 3' end of guanine nucleotides (GpN) and, like RNase A, produces oligonucleotides with terminal 3' phosphates.

Ribonucleic acid (RNA) A polymer of ribonucleoside monophosphates, synthesized by an RNA polymerase. RNA is the product of transcription.

Ribosomal RNA (rRNA) The predominant class of RNA in the cell. The highly abundant nature of rRNA makes it a useful indicator of the integrity, quality, and probable utility of a sample in downstream applications. The low complexity of this RNA species also makes it useful as a molecular weight marker for RNA electrophoresis.

Riboswitch An RNA genetic control element that experiences a conformational change upon metabolite (ligand) binding. Riboswitches can influence the cessation of transcription or initiation of translation.

Ribotyping An electrophoresis-based technique used to identify cells or microorganisms based on organism-specific variations in the structure of the ribosomal RNA genes.

Ribozyme An RNA molecule with the capacity to act as an enzyme. In other words, ribozymes are catalytic RNAs. Ribozymes may exhibit self-cleavage of their own backbone or may induce the cleavage of a complementary mRNA molecule to which the ribozyme base-pairs, presumably by folding into a characteristic three-dimensional shape.

RISC (RNA-induced silencing complex) A multiprotein endoribonuclease complex that mediates RNA interference by the cleavage of specific mRNA molecules to which a RISC-associated siRNA is complementary.

RLM-RACE (RNA ligase-mediated RACE) A form of RACE in which an RNA oligonucleotide is ligated to full-length mRNA molecules and then reverse transcribed into cDNA. This approach has the advantage of producing only clones that represent the precise 5' end of transcripts.

RNAi (RNA interference) Posttranscriptional gene silencing; occurs by coordinated destruction of specific mRNA molecules, thereby inhibiting translation of the associated protein. As the gene itself is not targeted or modified, the RNAi effect is reversible, thereby allowing functional studies of one gene at a time, especially related to developmental and disease pathways.

RNA polymerase A enzyme responsible for the synthesis of RNA polynucleotides by the process of transcription, using DNA as a template.

RT-PCR A method for the analysis of gene expression. mRNA is converted into cDNA by the enzyme reverse transcriptase and then amplified to the level of detection by PCR.

S1 nuclease An enzyme that degrades single-stranded nucleic acid molecules or any portion of a molecule, hybridized or otherwise, that remains in single-stranded form. The enzyme is zinc-dependent, requires an acidic pH, and has no known sequence preferences.

SAGE (serial analysis of gene expression) A bioinformatics tool for assessing the transcriptional activity of several genes under a defined set of experimental conditions by sequencing short cDNA molecules (tags) that have been randomly ligated together.

Saline sodium citrate (SSC) A salt solution frequently used for blotting of nucleic acids. It is also an essential component of various hybridization buffers and posthybridization washes. $20\times$ SSC = 3 M NaCl; 0.3 M Na₃-citrate; adjust pH to 7.0.

Saline sodium phosphate-EDTA (SSPE) A salt solution frequently used for blotting of

nucleic acids. It is also an essential component of various hybridization buffers and post-hybridization filter washes. The phosphate in this buffer mimics the phosphodiester backbone of nucleic acids, thereby providing enhanced blocking, lower background, and higher signal-to-noise ratio on membranes during Northern and Southern analysis. $20\times$ SSPE = 3 M NaCl, 0.2 M $\text{NaH}_2\text{PO}_4\cdot\text{H}_2\text{O}$, 0.02 M $\text{Na}_2\text{-EDTA}$; adjust pH to 7.4; autoclave.

Scorpions Fluorescent probes used in certain real-time PCR applications. Scorpions consist of an oligonucleotide primer linked directly to a fluorescent probe.

Sense RNA (also (+) RNA) Any RNA molecule that has the same nucleotide sequence as naturally occurring mRNA. Although *antisense* RNA is capable of hybridization with mRNA on a filter or *in situ*, sense RNA is only useful as a probe in Southern analysis; otherwise, sense RNA is generally used as a negative control to assess the magnitude of background (non-specific) hybridization.

Sevag Older term for a mixture of chloroform:isoamyl alcohol (24:1). *See*: Sevag, M.G., Lackman, D.B., and J. Smolens. (1938). The isolation of the components of streptococcal nucleoproteins in serologically active form. *J. Biol. Chem.* 124(2), 425–436.

shRNA (short hairpin RNA) A short RNA molecule with a characteristic hairpin shape that has a role in regulating gene expression via the RNA interference pathway.

siRNA (short interfering RNA) A short double-stranded RNA molecule with characteristic dinucleotide overhang at the 3' end of each strand. These molecules have a role in regulating gene expression via the RNA interference pathway.

Single nucleotide polymorphism (SNP) A single base change that is mostly commonly associated with a variant phenotype and, in some cases, with the onset or progression of a disease.

Slot blot A membrane-based technique for the quantitation of specific RNA or DNA

sequences in a sample. The sample is usually applied directly onto a filter by vacuum filtration through a manifold in either a slot or dot configuration (*see also* dot blot). Slot blots lack the qualitative component associated with electrophoresis-based methods, and are generally considered low sensitivity assays.

Sodium dodecyl sulfate (SDS) An ionic detergent commonly used to disrupt biological membranes and to inhibit RNase.

Southern analysis (also Southern blotting) A technique for transferring electrophoretically chromatographed DNA from an agarose gel matrix onto a filter paper for subsequent immobilization and hybridization. The information gained from Southern analysis is used to qualitatively and quantitatively assess the organization of specific genes or other loci.

Specific activity The amount of radioactivity per unit mass of a radioactive material. It is most frequently expressed in curies per millimole of material (Ci/mmol).

SSC *See* saline sodium citrate.

SSH (suppression subtractive hybridization) A method for identification of differentially expressed genes. Briefly, the mechanics of SSH result in a hybridization-based physical subtraction of sequences shared by two or more cDNA populations. The sequences which remain after the subtraction are PCR-amplified and represent those genes that were up- or downregulated under a particular set of conditions.

SSPE *See* saline sodium phosphate-EDTA.

Strand displacement amplification An isothermal alternative to PCR that is used for the exponential amplification of DNA.

Steady-state RNA The final accumulation of RNA in the cell or cytoplasm. For example, measurement of the steady-state abundance of a particular species of mRNA does not necessarily relate to the *rate* of transcription, nor about the age or turnover rate of specific RNAs. *Compare with* nuclear runoff assay.

Stoffel fragment (AmpliTaq[®] DNA polymerase) A thermostable recombinant DNA polymerase that is smaller (by 289 amino acids) than the full-length AmpliTaq polymerase. The Stoffel fragment is more thermostable, has activity over a broader range of Mg⁺⁺ concentrations and lacks a 5'→3' exonuclease activity. It is commonly used in multiplex PCR applications.

Stringency A measure of the likelihood that double-stranded nucleic acid molecules will dissociate into their constituent single strands; it is also a measure of the ability of single-stranded nucleic acid molecules to discriminate between other molecules that have a high degree of complementarity and those that have a low degree of complementarity. High-stringency conditions favor stable hybridization only between nucleic acid molecules with a high degree of complementarity. As stringency is lowered, a proportional increase in non-specific hybridization is favored.

Structural gene A gene that encodes an mRNA, as opposed to a rRNA or tRNA. Structural genes are usually single-copy or moderately repetitive genes, as demonstrated by C₀t kinetics.

Sumoylation Addition of a SUMO protein, which is similar to ubiquitin, to another protein. Sumoylation is a reversible modification. Unlike ubiquitination, which marks proteins for destruction, sumoylation regulates diverse cellular functions, including apoptosis, cell cycle control, stress-related cellular responses, and transcription regulation of gene expression, to name but a few.

SYBR[®] Green, SYBR[®] Gold Members of a new family of dyes for staining nucleic acids. Commonly provided as a 10,000× stock solution in DMSO, SYBR dyes are diluted to a working concentration of 1× in Tris buffer, such as 1× Tris-acetate-EDTA (TAE) buffer or 1× Tris-borate-EDTA (TBE) buffer. Among the advantages of using the SYBR dyes are greatly reduced background fluorescence, higher sensitivity, and reduced mutagenicity when compared with ethidium bromide. SYBR Green I

and SYBR Gold are used to stain DNA while SYBR Green II is used to stain RNA.

SYBR Green assay A platform for real-time PCR in which the dye SYBR Green I binds to double-stranded PCR products as they accumulate in the reaction tube.

Systems biology An emerging field of biology characterized by the systematic study and integration of biological information pertaining to DNA, RNA, proteins, signal transduction pathways, and the like in order to facilitate an understanding of the dynamic relationships between individual members of these various categories of biomolecules and networks, and how they influence both the cell and the organism.

TAE buffer See Tris-acetate-EDTA buffer.

TAP (tobacco acid pyrophosphatase) An enzyme that removes the 5' cap structure from eukaryotic mRNA. This enzyme is widely used in RLM-RACE.

TATAA box An AT-rich regulatory element that is associated with the promoter for eukaryotic RNA polymerase II. The TATAA box is usually found within nucleotides upstream from the actual transcription start site.

Taq DNA polymerase Thermostable, highly processive DNA polymerase from the organism *Thermus aquaticus*. Taq polymerase is one of several enzymes that can be used to support PCR.

TaqMan[®] assay A real-time PCR platform in which fluorescent probe is included in a reaction along with the primers. The accumulation of double-stranded PCR product leads to proportionate increase in fluorescence which is measured by the instrumentation.

Target Single-stranded DNA or RNA sequences that are complementary to a nucleic acid probe. Target sequences may be immobilized on a solid support, may be available for hybridization in solution, or may function as template molecules (see below) for amplification by PCR.

TBE Buffer See Tris-Borate-EDTA buffer.

TEMED (*N,N,N',N'*-Tetramethylethylenediamine) This quaternary amine is used as a catalyst in the polymerization of acrylamide in the preparation of polyacrylamide gels.

Template A macromolecular informational blueprint for the synthesis of another macromolecule. All polymerization reactions, including replication, transcription, and PCR, require templates, which dictate the precise order of nucleotides in the nascent strand. Primer extension-type reactions cannot proceed in the absence of template material.

Thermal cycler A programmable instrument in which PCR is performed.

Thymidine A nucleoside containing the base thymine.

Thymine A pyrimidine that usually base-pairs to adenine.

TIFF Tagged image file format; a type of software file used to store images.

T_m Melting temperature. That temperature at which 50% of all possible duplexes are dissociated into their constituent single strands. To facilitate formation of all possible duplexes, hybridization is conducted below the T_m of the duplex; the lower the temperature, the greater the likelihood that duplexes, including those with mismatches, will form.

TOPO[®] cloning An innovative method of ligating DNA via the enzyme topoisomerase. The method is much faster than traditional ligations involving T4 DNA ligase, and is compatible with blunt end- and sticky end cloning, as well as the cloning of PCR products with a 3' A-overhang.

Touchdown PCR An excellent method for preventing non-specific product formation. In touchdown PCR the initial annealing temperature (T_a) is *higher* than the T_m and is subsequently lowered in each passing cycle. Eventually the annealing temperature will drop below the T_m, thereby allowing amplification of only those template molecules which are exactly complementary to the primers.

Trailer sequence The untranslated sequence located just 3' to the coding region of mRNA. In poly(A)⁺ mRNA, the trailer sequence is located between the coding region and the beginning of the poly(A) tail.

Transcription The process by which RNA molecules are synthesized from a DNA template by RNA polymerase.

Transcription elongation The second of three stages of RNA synthesis, characterized by the covalent addition of nucleotides to the 3' end of the nascent chain. As the polymerase moves along the template, the newly synthesized RNA is displaced, permitting reannealing of the DNA template.

Transcription factor Any of a number of molecules needed to recruit RNA polymerase to a specific promoter and support related functions and, in so doing, support the initiation of transcription.

Transcription initiation The first of three stages of RNA synthesis. Initiation begins with the binding of RNA polymerase and associated transcription factors to double-stranded DNA, followed by the local unwinding of this template.

Transcription start site (TSS) The location of the first transcribed nucleotide near the 5' end of a gene.

Transcription termination The third and last stage of RNA synthesis. This involves recognition of the point beyond which no additional nucleotides are to be added. Following addition of the last base to the RNA chain, both the RNA and the RNA polymerase dissociate from the DNA template.

Transcriptome The complete set of mRNA molecules produced by a particular cell under a particular set of conditions.

Transcriptomics The all-inclusive analysis of an organism's transcripts and how the expression of mRNA molecules changes as conditions inside the cell change.

Transfection The introduction of nucleic acid molecules into a eukaryotic cell, most

often by chemical means or viral delivery, causing phenotype and genotype conversion of the recipient cell. Transfection is classified as either transient, meaning that there is eventually reversion of the cell to the pre-transfection phenotype, or stable transfection, in which the transfected DNA integrates into the chromosome of the host. Stable transfection of a eukaryotic cell is a rare event.

Transfer RNA (tRNA) A moderately abundant class of RNA molecules that shuttle amino acids to the aminoacyl site of the ribosome during protein synthesis. The total mass of tRNA in the cell is occasionally assayed to demonstrate that a particular experimental manipulation has not resulted in a change in overall transcription in the cell. While tRNA concentration is perfectly suitable as a gel loading control, it is not well-suited as a transcription control because tRNAs are produced in the cell by RNA polymerase III while mRNAs are produced by RNA polymerase II.

Translation The process by which peptides are synthesized from the instructions encoded within an RNA template. Translation occurs as mRNA is deciphered by the ribosomes.

Translation elongation The second of three stages of protein synthesis. Elongation encompasses all of the peptide bonds that are formed with the subsequent addition of amino acids to the carboxy terminus of the nascent polypeptide.

Translation initiation The first of three stages of protein synthesis. Initiation involves the association between mRNA and the ribosome, all of the biochemistry preceding this event, and the formation of the first peptide bond between the amino acids that will be located at the amino terminus of the nascent polypeptide.

Translation termination The third and last stage of protein synthesis, involving the release of both the completed polypeptide and the mRNA that directed its synthesis.

Triploid Having three complete sets of chromosomes (three of each chromosome). See also *diploid* and *haploid*.

Tris-acetate-EDTA buffer (TAE buffer) Common electrolyte reagent for the electrophoresis of DNA. $50\times$ TAE (per liter) = 242 g Tris base; 100 ml 0.5 M Na₂-EDTA, pH 8.0; 57.1 ml glacial acetic acid; autoclave or filter-sterilize. Working concentration is $1\times$ TAE.

Tris-borate-EDTA buffer (TBE buffer) Common electrolyte reagent for the electrophoresis of DNA, especially low molecular weight molecules. $20\times$ TBE (per liter) = 121 g Tris base; 61.7 g sodium borate; 7.44 g Na₂-EDTA; filter sterilize. Working concentration is $1\times$ TBE.

Ubiquitylation The process of tagging a protein for destruction by the addition of ubiquitin, a highly conserved 76-amino acid protein. This process is also known as ubiquitination.

Upstream Sequences in the 5' direction (away from the direction of expression) from some reference point. For example, the 5' cap in eukaryotic mRNA is located *upstream* from the initiation codon.

Uracil A pyrimidine base that usually base-pairs to adenine.

Uracil-N-glycosylase (UNG) An enzyme used to destroy dUTP-labeled PCR products from previous amplifications in order to eliminate carryover contamination in future reactions.

Uridine A nucleoside containing the base uracil.

Vanadyl ribonucleoside complexes (VDR; also VRC) An inhibitor of RNase that may be added to gentle RNA lysis buffers to control RNase activity. VDR functions as an RNA analog. The use of VDR has fallen sharply out of favor with most RNA efficians because even trace amounts of carryover VDR have been shown to be highly inhibitory toward reverse transcriptase, thereby making it unsuitable for RT-PCR applications.

Vector A nucleic acid molecule such as a plasmid, bacteriophage, or phagemid into which another nucleic acid molecule (the so-called insert or foreign DNA) has been ligated. Vectors contain sequences that, in a suitable host, permit propagation of the vector and the DNA that it carries.

Western analysis (also Western blotting) A technique for transferring electrophoretically chromatographed protein from a polyacrylamide gel matrix onto a filter paper for subsequent characterization by antigen-antibody recognition. The information gained from Western analysis is used to qualitatively and quantitatively assess the prevalence of specific polypeptides.

Whole genome amplification (WGA) Amplification of the entirety of a genome when very

limited amounts of starting material are available. Methods include fairly low-temperature annealing of 15-mer random primers or semi-degenerate oligonucleotides to ensure complete genome coverage. There are many forms of this assay which render a highly uniform, high fidelity whole genome amplification, and which eliminate the sequence bias associated with PCR-based methods. These systems are popular when working with archived, formalin-fixed tissue samples, as well as many other applications.

Xenobiotic A chemical that is foreign to a cell or organism.

Zwitterion An ion that carries both a positive and negative charge simultaneously.

Index

Note: Page numbers followed by f indicate a figure; page numbers followed by t indicate tabular material.

- A**
A₂₃₀ absorbance, 589
A₂₄₀ absorbance, 589
A₂₆₀ absorbance, 589
A₂₈₀ absorbance, 589
AAUAAA motif, 27, 123
Abbreviations, 665–666
Ab microarrays, 512
Absorption ratios, 144
Abundance categories, 122
Accession number, 563
Acridine orange, 210, 235
β-actin, 152
Adapters, 371
Adenine, 3f
Adenosine-5'-triphosphate, 5
Adipose tissue, RNA isolation from, 95–96
AE buffer, 597
Aerosol contamination, 397
Affinity chromatography, 126
Affinity matrices, in RNA, 73–74
Affymetrix, 512
Agarose
 gel staining techniques
 acridine orange, 210, 235
 ethidium bromide, 204–206
 GelStar, 209
 methylene blue, 210–211
 silver staining, 209–210
 SYBR Gold, 207
 SYBR Green, 207, 208t, 209f
 SYBR Safe, 208–209
 tips for first-time gel, 213–216
AG dinucleotide, 16
Alkaline blotting, 249–250
Alkaline phosphatase, 270, 313
α-amanitin, 18t, 419
β-mercaptoethanol, 142
Amino acid sequence, 400
Aminoallyl dCTP, 268
Aminoallyl dUTP, 268
Ammonium sulfate, 412t
Ampicillin, 597
AMV-RT. *See* Avian myeloblastosis virus reverse transcriptase
Anchorage-dependent cells, trypsinization protocol for, 619–620
Annealing of primers, 390
Anticodon
 defined, 9
 tRNA, 9
Antisense RNA
 defined, 541
 description of, 542
 probes, 276
 characteristics of, 278
 description of, 276–277
 3'-end-labeling of, 279
 5'-end-labeling of, 279
 hybridization, 276
 synthesis, 278–279
 in vitro transcription, 276f, 277f, 278–279
 therapeutic uses of, 542
Antisense strand, 38
AP. *See* Alkaline phosphatase
Arabidopsis thaliana, 11
ARE. *See* AU-rich element
Argonaute, 541
Aspergillus oryzae, 325
A-tailing, 438–439
ATGCAAAT, 13
ATTTGCAT, 13
AUG codons, 109

- AU-rich element, 28
- Autoclaving, 54
- Autoradiography, 263, 302–304
 applications of, 303–304
 cassette for, 309
 film
 development of, 310
 exposure time for, 306–307
 fixation, 310
 manual development of, 304
 preflashing, 309
 safelight for, 306
 X-ray, 305–306
 filter membranes, 304–305
 fluorography, 308–309
 intensifying screens, 307–308
 latent image, 302–303
 macroautoradiography, 302
 microautoradiography, 302
 protocol for, 310–312
- AUAAAA, 28
- Avian myeloblastosis virus reverse transcriptase, 373, 393
- Avidin, 126, 269, 313
- B**
- BAC. *See* Bacterial artificial chromosomes
- Bacterial artificial chromosomes, 564–565
- Bacteriophage promoters, 276
- Bacteriophage vectors, 382
- Bark, RNA isolation from, 112–113
- Basic Local Alignment Search Tool software, 563, 571–573
- BeadBeater, for disruption of cell, 87
- Beer's law, 142
- Betaine, 412t
- Bicistronic mRNA, 30–32
 prokaryotic mRNAs, 30–31
 in vitro applications, 30
- Bio-11-dUTP, 268, 269, 270f
- Bio-21-dUTP, 268
- Bioinformatics
 Basic Local Alignment Search Tool software, 563, 571–573
 data mining, 564
 defined, 563
 description of, 562, 571
 National Center for Biotechnology Information, 571
- BioMax film, 305
- Biotin
 detection using, 268–270, 313
 labeling uses of, 268–270, 270f
 nucleic acid probe labeling using, 268–270, 270f
 photoactivatable, 269
 psoralen, 269
 streptavidin, 269–270
- Biotinylated primers, 268–269
- Biotinylation, 268–269
- Bio-11-UTP, 268
- Bio-16-UTP, 268
- Bits-per-pixel, 225
- BLAST. *See* Basic Local Alignment Search Tool software
- Blocking. *See* Prehybridization
- β -ME. *See* β -mercaptoethanol
- β -mercaptoethanol, 57
- BMP files, 229
- Bovine serum albumin, 412t
- BPP. *See* Bits-per-pixel
- Brightness, 226
- BSA. *See* Bovine serum albumin
- Buffers
 AE, 597
 chaotropic lysis, 56–57
 common formulations of, 597–601
 denaturation, 597
 for formaldehyde gels, 187
 gentle lysis, 49–51
 guanidinium, 57–58
 hybridization, 598
 loading, 213, 215, 598
 lysis, 56–57
 MOPS, 598
 neutralization, 599
 phosphate, 599
 TE, 600
 TE-9, 600
 tris/SDS, 601
- Burkitt's lymphoma, 14
- N-butanol, 269
- C**
- CAAT box, 13
- Caenorhabditis elegans*, 539

- Calcium tungstate phosphor-coated screens, 307–308
- Calf intestinal phosphatase, 426
- Cancer panel, 502
- Canonical base pairs, 8
- Capillary transfer, for Northern analysis, 246–247, 250–252
- Capping, of eukaryotic mRNA, 23
- CaWO₄-coated screens. *See* Calcium tungstate phosphor-coated screens
- CCD camera. *See* Charge-coupled device camera
- cDNA, 2, 386
 - anchor, 427
 - first-strand synthesis, 369–372, 377–379
 - libraries, 371
 - microarrays, 501
 - second-strand, 375–376, 377, 379
 - synthesis, 451, 452, 581
 - adapter-mediated double-stranded, 380
 - advantage of, 368
 - classical methods, 376–377
 - cloning procedures, 349
 - cloning systems, 381
 - dATP, dCTP, dGTP, amino allyl-dUTP, 508
 - efficiency of, 451
 - electrophoresis, first-strand assessment, 380
 - enzymes of, 369–370
 - gene expression, 383
 - Gubler and Hoffman method, 377
 - mechanics of, 368
 - Northern blot analysis, 391
 - PCR amplification, 328
 - PCR-based methods, 377
 - reverse transcriptase options, 372–375
 - reverse transcription, efficiency of, 379
 - SMART™ technology, template-switching, 375
- cDNA library
 - construction of, 382
 - creation of, 368
 - definition of, 368
 - microarray *vs.*, 504
 - older, 581
 - oligo(dT)-primed, 371–372, 581
 - poly(A)⁺ mRNA, 124–125
 - screening, 581–582
- CDP Star, 315
- Cell culture, advantages of, 82–83
- Cell freezing ampoules, 347
- Cell lysis, 46
- Cellular RNA, 494
- Central dogma, 10, 10f
 - of molecular biology, 10–11
- Centricon devices, 280
- Centrifugation
 - applications, 651–652
 - definition, 649
 - density gradient centrifugation
 - isopycnic technique, 653–654
 - sedimentation velocity, 652–653
 - description of, 589
 - differential, 652
 - rotors, 650–651
- Centrifuges, 649–650
- Centrifuge tubes, 617
- Cesium chloride, 167–168
 - density gradient centrifugation, 61–65
 - description of, 57
- Cesium trifluoroacetate, 57, 168
 - advance preparation of, 66–68
 - density gradient centrifugation, 65–68
 - disadvantages, 65
 - protocols for, 65–66
- Cetyltrimethylammonium bromide, 108
- Chaotropic lysis buffers, 56–57, 155
- Chargaff's Rule, 5
- Charge-coupled device camera, 232
- Chemiluminescence detection, 312
 - alkaline phosphatase systems, 315
 - chromogenic procedure, 318
 - defined, 314
 - difficulties associated with, 317
 - probe molecules, 314–315
 - steps involved in, 316–317
 - substrates for, 315–317
 - vs.* autoradiography, 315t
- Chloroform addition to phenol reagents, 603–604
- Chloroplasts, 106
- Chloroplast transcripts, 199, 200f

- Chromogenic detection
 description, 312
 procedures for, 317–318
- CIP. *See* Calf intestinal phosphatase
- Cloning
 cDNA, 379–383
 DNA ligase and topoisomerase, 383
 ligation efficiency, 381–382
 directional, 371, 380
 of double-stranded DNA, 14
 TA
 description, 437
 ligation reaction, 439–440
 schematic diagram of, 438f
 TOPO, 440–441
- C-myc*, 284, 576
- Coding region, of mRNA, 25–26
- Coding strand, 14
- Codons
 AUG, 25–26
 for degenerate primer, 402t
- Competitive PCR
 competitor for
 non-homologous, 480–481
 quantity of, 474
 reverse transcription efficiency of, 476
 target template and, 475–476
 composite primers, 469, 470f
 considerations, 472–477
 defined, 472–473
 description of, 468
 detection method, 477
 ethidium bromide for, 477
 forms of, 468–469, 471
 heteroduplexes, 476
 image analysis, 485–486
 primary, 478, 479f
 primary amplification, 482–484
 protocol, 480–485
 secondary, 478, 479f
 secondary amplification, 484–485
 steps involved in, 477–479
 summary of, 477
 SYBR Green or Gold, 485
 target template
 competitor and, 475–476
 reverse transcription efficiency of, 476
 transcript quantification by, 473f
 troubleshooting of, 486–488
- Complementary DNA. *See* cDNA
- Contig, 563
- Continuous probe labeling, 264f
- Contrast, 226
- Conversion, mass to moles, 593–595
- cRNA amplification, 467
- CsCl. *See* Cesium chloride
- CSPD, 315
- CsTFA. *See* Cesium trifluoroacetate
- CTAB. *See* Cetyltrimethylammonium bromide
- Cy3 (indocarbocyanine), 267–268
- Cy5 (indodicarbocyanine), 267–268
- Cy3-labeled cDNA, 502
- Cy-labeled probes, purification of, 508
- Cytoplasmic RNA
 cellular, 578
 Northern analysis of, 575
 total, 578
- Cytosine, 3f
- D**
- Data mining, 564
- Degenerate primer design
 codon utilization, 402
 IUB base abbreviations, 403
- Deionization, 615
- Denaturation
 of probe, 294–295
 of RNA
 description of, 185–187
 formaldehyde, 186–191
 formamide, 187
 glyoxal/dimethyl sulfoxide, 191–194
 methylmercuric bromide, 187
 selection of, 186
 urea, 187, 189–191
- Denaturation buffer, 597
- Denhardt's solution, 290, 597
- Density gradient centrifugation
 cesium chloride for, 61–65
 cesium trifluoroacetate for, 65–68
 definition of, 60–61
- Density gradient centrifugation–isopycnic technique, 653–654
- Density gradient centrifugation–sedimentation velocity, 652–653

- Deoxyribonuclease, 166
DEPC. *See* Diethyl pyrocarbonate
Deproteinization, 48
Detection
 autoradiography. *See*
 Autoradiography
 chemiluminescence, 312
 alkaline phosphatase systems,
 315
 chromogenic procedure, 318
 defined, 314
 difficulties associated with,
 317
 probe molecules, 314–315
 steps involved in, 316–317
 substrates for, 315–317
 vs. autoradiography, 315t
 chromogenic
 description, 312
 procedures for, 317–318
 digital imaging systems for,
 318–320
 non-isotopic methods of, 312
 biotin, 313
 chemiluminescence, 314–317
 chromogenic, 317–318
 digoxigenin, 313
 direct enzyme labeling, 314
 fluorescence imaging, 314
 PhosphorImager, 319–320
 rationale for, 301
 Dextran sulfate, 290
DHFR. *See* Dihydrofolate reductase
Dicer, 541
Diethyl pyrocarbonate, 162–164
Differential centrifugation, 652
Differential display PCR
 advantages, 533, 533t
 background, 526–529
 cDNA downstream primer design,
 529–530
 cDNA synthesis in, 530
 disadvantages, 533t
 downstream primer in, 529–530
 gel loading strategy for, 527t
 PCR
 amplification of, 530–531
 RNA isolation for, 529
 steps involved in, 529–533
 troubleshooting, 534–536
Differential expression, 516–517
DIG. *See* Digoxigenin
DIG-11-dUTP, 270, 271f
Digital imaging
 description of, 221–222
 detection uses of, 318–320
 rationale for, 221–222
Digitalis, 270
Digital systems, for photodocumentation
 description of, 223–228, 233
 digitization process, 224
 example of, 223, 223f
 image analysis systems used with
 advantages of, 229–230
 background correction using,
 230
 benefits of, 229t
 budget for, 231–232
 image formats, 228–229
 misconceptions regarding, 230
 optical density, 228
 phosphorImagers, 232
 practical considerations for,
 229–231
 software used with, 223, 224f
 terminology associated with,
 224–228
 workshops, 232
Digoxigenin
 detection using, 313
 probe labeling using, 270–271
DIG-11-UTP, 270
Dihydrofolate reductase, 419
Dimethyl sulfoxide, 412t
 glyoxal and
 RNA denaturation using, 191–194,
 254–255
 tips for using, 586
Direct enzyme labeling
 detection using, 314
 of DNA probes, 271–272
Direct exposure, defined, 302
Directional cloning, 371, 380
Disposable generators, for tissue
 homogenization, 84–85
Dithiothreitol, 57
D-Max, 239
D-Min, 239

- DMSO. *See* Dimethyl sulfoxide
- DNA
- DNase I removal of, from RNA sample, 611–612
 - exons, 16
 - glyoxalated, 192
 - heterologous, 597
 - introns, 16
 - nucleotides, 4
 - RNA *vs.*, 4
 - simultaneous isolation of, 68–71
 - template of, 19, 22
- DNA array technology, 504
- DNA cloning techniques, 12
- DNA-directed RNA interference, defined, 541
- DNA end-labeling, 264f
- DNA isolation
- high-molecular-weight, by salting-out, 621–623
- DNA libraries, 368
- DNA ligase, 383
- DNA methylation, 495
- DNA modifications, 495
- DNA polymerase, 394
- description of, 4, 5
 - E. coli*, 376
 - RNA-dependent, 372
- DNA polymerases, I, 611–612
- DNA probes, 272
- double-stranded, 272
 - 3'-end-labeling of, 275
 - 5'-end-labeling of, 275
 - first-strand cDNA, 272
 - labeling of, 273–275
 - nick translation, 274
 - oligonucleotides, 273
 - PCR of, 274
 - random priming, 274
 - synthesis, 273–275
- DNase. *See* Deoxyribonuclease
- DNA sequences, 343, 353, 382, 501
- dNTP, 4
- Dot blot analysis
- advantages, 643–644
 - apparatus, 642f
 - data limitations, 648
 - disadvantages, 644
- DNA, 646–647
- negative controls for, 644–645
 - positive controls for, 644–645
 - RNA, 645–646
- Double-stranded DNA
- cloning of, 14
 - stability of, 174
- Double-stranded molecules, 7
- formation of, 173–174
 - stability of, 174
 - types of, 173–174
- Double-stranded RNA, 36
- defined, 541
 - formation of, 174
- Dounce homogenizer, 86–87
- Downstream, 7
- Driver, 516
- Drosha, defined, 541
- dsDNA. *See* Single-stranded DNA
- DTT. *See* Dithiothreitol
- Dual-channel array, 502
- Duplex formation, 283
- DuPont's Cronex Quanta III Intensifying Screen, 307
- Dynabeads, 127
- E**
- Eberwine linear amplification process, 442
- ECL system. *See* Enhanced Chemiluminescence system
- ECM. *See* Extracellular matrix
- EDTA. *See* Ethylenediaminetetraacetic acid
- 18S rRNA, 198, 199
- 8-hydroxyquinoline, 167
- Electroblotting, 249
- Electrophoresis
- agarose
 - concentration, 630
 - purity of, 630 - applications of, 625
 - applied voltage for, 631–632
 - base composition and temperature, 633
 - casting tray for, 213
 - considerations for, 213–216

- definition of, 625
 - electrocution concerns, 180, 216
 - equipment for, 212–213, 635–636
 - ethidium bromide, 632
 - field direction, 633
 - gel
 - boilover concerns, 214
 - casting of, 214
 - stains, 204–211
 - submersion of, 214–215
 - thickness of, 214
 - gel boxes, 212–213, 215, 633–634, 634f
 - loading buffer for, 213, 215
 - matrix for, 628–629
 - molecular size range, 631
 - molecular weight standards for RNA.
 - See Molecular weight standards, for electrophoresis
 - Northern analysis and, 180
 - nucleic acid conformation, 631
 - nucleic acid normalization for
 - description of, 180
 - issues associated with, 180–181
 - poly(A)⁺, 183–185
 - poly(T) probe for, 183–185
 - procedure for, 181–183
 - plant RNA, 199, 200f
 - polyacrylamide gels
 - concentration of, 630–631
 - description, 628
 - power supply for, 213
 - principles of, 625
 - rationale of, 179–180
 - ribonuclease contamination,
 - 213
 - of RNA, 192–194
 - RNA denaturing systems for
 - description of, 185–187
 - formaldehyde, 186–191
 - formamide, 187
 - glyoxal/dimethyl sulfoxide, 191–194
 - methylmercuric bromide, 187
 - selection of, 186
 - urea, 187, 189–191
 - running buffer for, 213
 - safety during, 212–213, 635
 - salt effects on, 216f
 - sample
 - matrix for, 628–629
 - volumes for, 215
 - sensitivity of, 182
 - sodium phosphate buffer for, 192
 - SYBR Green for, 632–633
 - terminology associated with, 213
 - theoretical considerations for, 626–628
 - versatility of, 180
 - well appearance, 201
- Electrophoretograms, 235–238
- Elutip-d minicolumns, 280
- Endogenous mRNA, 452
- Enhanced Chemiluminescence (ECL) system, 314
- Enhancer sequences, 14
- Enzyme blends, for RT-PCR, 411–412
- Enzyme DNA polymerase, 394
- Enzyme(s)
 - for cDNA synthesis, 369, 370t
 - for PCR, 411–412, 413t
- Enzyme uracil-N-glycosylase, 395
- Epi-illumination, 234
- Equipment suppliers, 655–662
- Escherichia coli* single-strand binding protein, 411
- EST. See Expressed Sequence Tag
- EtBr. See Ethidium bromide
- EtBr- or SYBR green-stained RNA gel, 148
- Ethidium bromide
 - competitive PCR use of, 477
 - description of, 204–206, 598
 - disposal of, 607–609
 - electrophoresis use of, 632
 - SYBR Green *versus*, 208t
- Ethidium bromide-stained gels, 485
- Ethylenediaminetetraacetic acid
 - description of, 589
 - stock solutions, 597
 - vanadyl ribonucleoside complexes
 - removal by, 159
- Eukaryotic cell, 345
- mRNA
 - capping of, 23
 - description of, 9
- RNA polymerases, 18–22
- transcription in, 20t

- Eukaryotic mRNA
 topology of, 23
- Eukaryotic RNA
 subpopulations of, 47t
- Exact-match probes, 288
- Exon-intron junctions, 500
- Exon(s)
 in cloned DNA, 17
 description of, 16
- Exon trapping, 16
- Expressed Sequence Tag, 563
- Extracellular matrix
 in vivo, 82
- F**
- Farrell's RNA sensitivity index, 450
- FastPrep system, 87
- FFPE. *See* Formalin-fixed
 paraffin-embedded
- Film
 preflashing, 309
 X-ray
 autoradiography uses of, 305–306,
 305t
 developing scheme for, 303f
 exposure time for, 306–307
 safelight for, 306
 development of, 236–237
 photodocumentation uses of,
 238–239
- Filter membranes, for Northern analysis,
 304–305
 description of, 242
 formaldehyde denaturation of, 254
 handling of, 245, 254–255,
 257–258
 nitrocellulose, 243–244
 nylon
 description of, 244–245
 ultraviolet cross linking RNA to,
 257
 postfixation handling of,
 257–258
 posttransfer handling of,
 254–255
 preparation of, 245–246
 purchase of, 242
 PVDF, 245
- First-strand cDNA synthesis, 373
- AMV RT, 377
 assessment of, 380
 mechanics of, 375
 PCR, 377
 priming mRNAs, 426
 RNA:DNA hybrid, 393
- 5' cap, mRNA, 23–25
- 5'-end-labeling
 of DNA probes, 264, 264f, 275
 of RNA probes, 279
- 5' nuclease assay for real-time PCR, 459
- 5S rRNA, 198
- 5'-5' triphosphate, 24
- 5' T7 RNA polymerase promoter, 466
- 5' untranslated region, 25
- Fixed tissue, RNA isolation from,
 91–92
- FL-12-dUTP, 271, 272f
- Fluor, defined, 308
- Fluorescein, 271
- Fluorescence, 314
- Fluorescent dyes, 508
- Fluorescent nucleotides, 268, 271
- Fluorogenic probes, 459
- Fluorography, 308–309
- FL-12-UTP, 271
- Formaldehyde
 buffer for, 187
 deionization, 615
 Northern analysis filter denaturation
 using, 254
 RNA denaturation using, 186–191
 toxicity of, 187
- Formalin-fixed paraffin-embedded, 91
- Formamide, 79, 412t
 deionization, 615
 melting temperature affected by, 175
 stringency affected by, 177
 stringency and, 175, 177
- Formamide, nucleic acid hybridization
 and, 290
- FORMAzol, 79
- Formyl-methionine, 25
- Forward kinasing reaction, 275, 279
- Forward subtraction, 516
- FOTODYNE Luminary/FX workstation,
 223
- 4-chloro-1-naphthol, 318
- Fresh tissue, RNA isolation from, 89–90

- Frozen tissue, RNA isolation from, 90–91
- Fruit, RNA isolation from, 113–115
- Functional genomics
 aptamers, 494
 bioinformatics, 494–495
 classical molecular biology, 496–497
 definition, 491–492
 epigenetic, 495
 gene/protein interactions, 492
 gene transfer, 494
 goals of, 491
 importance of, 492–493
 mass spectrometry, 495
 metabolomics, 495–496
 microarray analysis, 493
 nanobiotechnology, 494
 PCR, 494
 protein chips, 495
 RNA interference, 494
 serial analysis of gene expression, 494
 tools, 492
 two-dimensional gel electrophoresis, 495
- G**
- Gamma correction curve, 226, 227f
- GAPDH. *See* Glyceraldehyde-3-phosphate dehydrogenase
- Gas-liquid chromatography, 496
- G-box element, 13
- GBX-2 safe light filter, 306
- GC box, 13
- G+C content, 288
- Gel boxes, 212–213, 215
- Gel filtration, 410
- Gel-Pro Analyzer
 image analysis using, 223, 224f, 319
- Gel staining
 acridine orange, 210, 235
 ethidium bromide
 competitive PCR use of, 477
 description of, 204–206, 598
 disposal of, 607–609
 electrophoresis use of, 632
 SYBR Green *versus*, 208t
 GelStar, 209
 methylene blue, 210–211
 silver staining, 209–210
 SYBR Gold, 207
 SYBR Green
 description of, 207, 208t, 209f
 SYBR Safe, 208–209
 GelStar, 209
 GenBank, 261
 Gene expression, 422, 503
 description, 576
 microarray-based profiling, 504
 Gene rearrangement, 576
 Gene regulation, levels of, 32–34
 Gene(s)
 exons associated with, 16
 housekeeping, 181
 introns associated with, 16
 organization of, 14–18
 Gene silencing
 defined, 542
 posttranscriptional, 542
 RNA interference methods for, 544
 Gene-specific primers, 429, 433
 Genome, 564–565
 defined, 563
 Genomic DNA, amplification of, 394
 Genomics
 chemical, 563
 defined, 563
 description of, 564–565
 functional, 563
 Gentle hypotonic lysis, cytoplasmic RNA, 51–56
 isolation of, 53–56
 preparation of extraction buffer, 51–53
 procedure for, 51
 Gentle lysis buffers, 49–51
 disadvantage of, 50–51
 inclusion of non-ionic detergents, 49–50
 G-HCl. *See* Guanidine hydrochloride
 GIF files, 229
 Glass microarrays, 501
 GLC. *See* Gas-liquid chromatography
 Globin mRNA, 466
 Glyceraldehyde-3-phosphate dehydrogenase, 418
 Glycerol, 412t
 Glycerol-3-phosphate dehydrogenase, 434

- Glyoxal, deionization, 615
 G3PD. *See* Glycerol-3-phosphate dehydrogenase
 Gram-negative bacteria, 15
 Gram-positive bacteria, 15
 Gray-scale pixels, 255–256
 GTC. *See* Guanidine thiocyanate
 Guanidine hydrochloride, 165
 Guanidine thiocyanate, 165–166, 598
 Guanidinium-acid-phenol extraction, 58–60, 534
 protocol for, 59–60
 Guanidinium buffers
 RNA following cellular disruption, 58f
 RNA isolation with, 57–58
 Guanidinium isothiocyanate, 437
 Guanidinium thiocyanate, 59
 Guanine, 3f, 406
 Guanosine nucleotide, 24
 GU dinucleotide, 16
 GUG, 25
- H**
³H, 265, 265t, 266
 Half-life
 biological, 266
 radioactive, 266
 Hand-held motorized homogenizer, 85f
 HCONH₂. *See* Formamide
 Heat map with hierarchical clustering of genes, 505
 Heteroduplexes, 476
 Heterogeneous nuclear ribonucleoprotein complexes, 17
 Heterogenous nuclear RNA, 8, 17, 57, 179
 5' cap, 23
 splicing of, 33
 HGPR. *See* Hypoxanthine-guanosine phosphoribosyltransferase
 High-molecular-weight DNA isolation, by salting-out, 621–623
 High performance liquid chromatography, 410, 496
 High-throughput screening
 microarrays for. *See* Microarrays
 overview, 499
 Histones, 418
 HIV reverse transcriptase, 373
 hnRNA. *See* Heterogenous nuclear RNA
 hnRNP. *See* Heterogeneous nuclear ribonucleoprotein complexes
 H₂O₂. *See* Hydrogen peroxide
 Hogness box. *See* TATA boxes
 Homogenization
 description of, 84
 Dounce homogenizer, 86–87
 motorized devices for, 84
 polytron disruption, 84–86
Homo sapiens, 160
 Horseradish peroxidase, 313
 Housekeeping genes, 181, 456
 normalization, 417
 HPLC. *See* High performance liquid chromatography
 HRP. *See* Horseradish peroxidase
 Human cells, rRNA biogenesis in, 19
 Human immunodeficiency virus, 373
 Hybridization
 definition of, 261
 detection, by chemiluminescence, 316
 mixed-phase, 261
 nucleic acid. *See* Nucleic acid hybridization
 of nucleic acid molecules, 175
 RNA probes for, 276
 in situ, 90
 solution, 575, 579
 subtraction, 515–516
 in subtraction-suppression PCR, 520–521
 variables that affect, 175
 Hybridization buffer, 598
 Hydrogen bonding, 6, 126
 Hydrogen peroxide, 164, 316
 Hydroxyl group (-OH), 3
 Hyperchromic shift, 286
 Hyperfilm, 305
 Hypotonic lysis
 cytoplasmic RNA isolation using, 51–56
 nuclear isolation using, 347
 Hypoxanthine-guanosine phosphoribosyltransferase, 418

- I**
- Image analysis systems
 advantages of, 229–230
 background correction using, 230
 benefits of, 229t
 budget for, 231–232
 image formats, 228–229
 misconceptions regarding, 230
 optical density, 228
 phosphorImagers, 232
 practical considerations for, 229–231
 software used with, 223, 224f
 terminology associated with, 224–228
 workshops, 232
- Informosomes, 32
- In-gel autoradiography, 360
- Initiation codon, 23f, 25
- In silico* proteomics, 562
- In situ* autoradiography. *See*
 Microautoradiography
- Integrated optical density, 223, 228
- Intensifying screens, 307–308
- Intercalator-based detection, 457
- Interference RNA. *See* RNA interference
- Internal ribosome entry site (IRES), 30
- Internet resources, 583
- Introns
 description of, 16
 length of, 16
 removal of, 17
 splicing of, 16–17
 termination codon, 16
- In vitro* transcription, 276f, 277f,
 278–279
- In vitro* transcription
 template construction, 469
- In vitro* translation
 of mRNA, 2
- IOD. *See* Integrated optical density
- Ionic strength
 nucleic acid hybridization and,
 286–287
- IRES. *See* Internal ribosome entry site
- Isoamyl alcohol addition to phenol
 reagents, 603
- Isolation
 of nuclear RNA, 361–362
 of polysomal mRNA, 98–99
 of RNA. *See* RNA isolation
- Isopycnic centrifugation
 cesium chloride, 61–65
 defined, 60–61
 density gradient, 653–654
- Isotopic labeling, of nucleic acid probes,
 265–267
 decomposition problems, 266–267
 disadvantages of, 266
 ³H, 265, 265t, 266
 half-life of, 265
 ³²P, 265t, 266, 267
 ³³P, 265–266, 265t
 ³⁵S, 265, 265t, 266
 safety considerations, 265
- J**
- JPEG files, 229
- K**
- Kinasing reaction, 264, 264f. *See also*
 5'-end-labeling, of DNA
 probes
- Kits
 for cDNA synthesis, 369
 for RNA isolation, 71–74
 for RT-PCR optimization, 411
- Klenow fragment of DNA polymerase I,
 274
- Kodak D-19 and D-76, 304
- L**
- Labeling, of nucleic acid probes
 isotopic, 265–267
 decomposition problems, 266–267
 disadvantages of, 266
 ³H, 265, 265t, 266
 ³²P, 265t, 266, 267
 ³³P, 265–266, 265t
 ³⁵S, 265, 265t, 266
 safety considerations, 265
 nonisotopic
 biotin, 268–270
 Cy3, 267–268
 Cy5, 267–268
 digoxigenin, 270–271
 direct enzyme labeling, 271–272
 fluorescein, 271
 methods of, 267
- Laboratory records, 591–592

- Lac* operon, 15
Latent image
 defined, 302
 formation of, 302–303
Latent ribonucleases, 157–158, 158f
Lauria-Bertani Medium, 598
Lauryl sulfate, 166
LB Medium. *See* Lauria-Bertani Medium
LCR. *See* Ligase chain reaction
Leader sequence, 25
Leaf, RNA isolation from, 110–112
LiCl-Urea method, RNA isolation from tissue, 92–95
Ligase chain reaction, 443
Linear RNA amplification, 441
Lipid-enriched, RNA isolation from, 95–96
LNAs. *See* Locked nucleic acids
Loading buffer, 213, 215, 598
Locked nucleic acids, 494
 RNA/DNA oligonucleotides, 415
Long probes
 melting temperature for, 291
 prehybridization for, 293–294
Long-range PCR, 429–431
Low abundance mRNA, 22, 122
Lumigen APS-5, 315–316
Lumigen PPD, 315
Luminol, 316, 317f
Lumi-Phos 530, 315
Lumi-Phos Plus, 315
LUX technology, 460
Lyse cells, 84
Lysis buffer
 chaotropic, 56–57
 formulations of, 49
 gentle, 49–56
 guanidinium, 57–58
- M**
Macroautoradiography, 302
Magnetic bead
 poly (A)⁺ mRNA purification by, 127–129
Mass determination
 cDNA, 484
 primary tenfold dilution series, 479
 secondary twofold dilution series, 479
Mass standard, defined, 145
Mass to moles, conversion of, 593–595
Material safety data sheet, 73, 590
MCS. *See* Multiple cloning site
Melting temperature
 description, 285–286
 formamide effects on, 175, 290
 formula, 406
 for long probes, 291
 for oligonucleotide probes, 292
 PCR primers, 406
Mendel, Gregor, 10
Messenger ribonucleoprotein complexes, 27, 185
Messenger RNA, 22–32
 affects in translation, 31–32
 biogenesis, 342
 denaturation, 378
 in eukaryotic cells, 22
 maturation, 342
 purification of, 125t
 splicing of, 34
 stability of, 29–30
 template, 375, 398
 topology of
 5' cap, 23–25
 coding region, 25–26
 organellar, 28–29
 poly(A) tail, 26–28
 trailer sequence, 26
 transport of, 29–30
 trans-splicing of, 35–36
 turnover of, 29–30
 in vitro transcription of, 2
 in vitro translation of, 2
Metabolomics, 563
Methylation, 24
Methylene blue, 210–211
Methylmercuric bromide, 187
M⁷G. *See* 7-methylguanosine
Microarray analysis, 500
 basic steps, 507–509
 cDNA libraries, 504
 gene expression, 511
 scatter plot analysis, 509
SSH, difference between, 506
steps, 503

- Microarray-based assays, 499
- Microarray-based profiling, advantage of, 504
- Microarrays
- Ab, 512
 - advantages, 511
 - analysis of. *See* Microarray analysis
 - antibody, 500
 - categories, 500
 - cDNA/cRNA, 502
 - cDNA libraries, 504
 - contemporary, 500
 - disadvantages, 511
 - gene sequences on, 507
 - glass, 501
 - multiple-tissue, 502
 - nylon, 493, 501
 - pairwise design, 510
 - printing of, 501
 - qualitative information, 507
 - quality control, 501
 - real-time PCR *vs.* suppression subtractive hybridization methods, 506
 - reference RNA design, 510
 - schematic diagram of, 503f
 - semiquantitative nature of, 506
 - short oligonucleotides, 501
 - synthesis of, 501
 - themes of, 502
- Microautoradiography, 302
- Microcentrifuges, 649
- Microfuge tubes, 161
- microRNA, 26, 359, 549–550
- defined, 541
 - regulation of gene expression, 37–38
 - RISC-associated with, 38
- Minor groove binder, 458
- miRNA. *See* microRNA
- Miscellaneous lab tips, 585–590
- Mismatching, nucleic acid hybridization and, 288–289
- Mixed-phase hybridization
- definition of, 261
- MMLV-RT. *See* Moloney murine leukemia virus
- Molecular beacons, 460
- Molecular biology
- rules of, 585
 - tips for, 585–590
- Molecular weight standards, for electrophoresis
- description of, 194–195
 - external, 195
 - internal, 195
 - range of, 197
 - reliability of, 203–204
 - ribosomal RNA as, 197–204
 - use of, 195–197
- Moloney murine leukemia virus
- reverse transcriptase, 373–374
- Monochrome, 225
- Monocistronic, 15
- MOPS buffer, 598
- mRNA differential display, 496
- advantages, 533, 533t
 - background, 526–529
 - cDNA downstream primer design, 529–530
 - cDNA synthesis in, 530
 - disadvantages, 533t
 - downstream primer in, 529–530
 - gel loading strategy for, 527t
 - PCR
 - amplification of, 530–531
 - RNA isolation for, 529
 - steps involved in, 529–533
 - troubleshooting, 534–536
- mRNP. *See* Messenger ribonucleoprotein
- mRNP complexes. *See* Messenger ribonucleoprotein complexes
- MSDS. *See* Material safety data sheet
- Multiple cDNA syntheses, 435
- Multiple cloning site, 438
- Multiple-tissue microarrays, 502
- Multiplex primer design, 407
- Mycoplasma* contamination, 507
- N**
- Nanobiotechnology, 494
- NanoDrop, 142
- spectrophotometer, 143f
- NanoDrop spectrophotometric analysis, 380
- NaOH. *See* Sodium hydroxide
- NASBA. *See* Nucleic acid sequence based amplification

- National Center for Biotechnology Information, 495, 571
- NCBI. *See* National Center for Biotechnology Information
- ncRNA. *See* Non-coding RNA
- Negative control reactions, 471
- Nested PCR, 428–429, 428f
- Neutralization buffer, 599
- Nick translation, DNA probe labeling and, 274
- Nitrocellulose, 243–244
- N*-laurylsarcosine, 148, 166
- Non-coding RNA, 37
- Non-denaturing gels
RNA on, 194
- Non-homologous competitor DNA template, 470
- Nonidet P, 40, 46
preparation of, 599
- Non-ionic detergents, 412t
- Nonspecific inhibitors, of ribonucleases, 160–161
- Non-translated leader sequence, 25
- Northern analysis, 261, 262
cytoplasmic RNA, 575
definition of, 242
description, 578–579
electrophoresis and, 180
filter membranes for
description of, 242
formaldehyde denaturation of, 254
handling of, 245, 254–255
nitrocellulose, 243–244
nylon, 244–245
posttransfer handling of, 254–255
preparation of, 245–246
purchase of, 242
PVDF, 245
hybridization and, 292–296
nucleic acid immobilization techniques
baking, 255–256
cross linking by ultraviolet irradiation, 256–257
description of, 255
quality control technique, 153
rationale for, 241–242
reverse
advantage of, 259
convenience of, 259
description of, 258, 259f
schematic diagram of, 578f
sensitivity of, 182
transfer techniques
alkaline blotting, 249–250
capillary, 246–247, 250–252
description of, 246
electroblotting, 249
passive capillary, 246–247, 250–252
TurboBlotter, 247–248, 248f, 252–254
vacuum blotting, 248–249
- NP-40. *See* Nonidet P 40
- Nuclear RNA
difficulties, 363–364
drawback of, 361
heterogenous
description of, 179
isolation of, 361–362
northern analysis of, 360
preparation of, 362–363
pulse labeling tissue culture cells, 357
steady-state analysis, extraction of, 360–361
transcriptional efficiency, 341
transcriptional elongation, 343
- Nuclear run-off assay, 343, 346
advantage of, 343, 345
cells, harvesting of, 349
cumbersome/labor-intensive, 356
fragile nuclei, 349–350
labeled RNA transcripts, 356–357
labeling reaction, 348
nuclei, preparation of, 349–351
posthybridization washes and detection, 355
RNA, preparation of, 354–355
target DNA, preparation of, 353–354
transcripts, labeling/recovery of, 351–353
- Nuclear runoff assay, description of, 579f
- Nuclease protection assay, 302, 321
approach to, 322–327
background associated with, 330–331
bands in, 332
description of, 579f
double-stranded probes used with, 323

- lack of signal, 331–332
- Northern analysis and, 328–329, 328f
- optimization of, 329–330
- potential difficulties, 330–332
- probes for
 - selection of, 327–329
 - single-stranded, 327
- RNase
 - characteristics of, 325
 - protocol for, 335–337
 - RNase A and RNase T1, comparison of, 326t
 - RNA transcript quantification using, 324f
 - S1 nuclease assay *vs.*, 326t
- S1 nuclease assay
 - characteristics of, 325
 - disadvantages of, 325
 - protocol for, 332–334
 - RNase *vs.*, 326t
 - RNA transcript quantification using, 323f
- troubleshooting, 337
- Nuclease protection–pulse label transcription assay, 357–359
- Nucleic acid(s)
 - glyoxalated, 192
 - immobilization of, for Northern analysis
 - baking, 255–256
 - cross linking by ultraviolet irradiation, 256–257
 - description of, 255
 - migration of, in electrical fields, 186
 - non-spectrophotometric methods for determining, 145–147
 - normalization of, for electrophoresis
 - description of, 180
 - issues associated with, 180–181
 - poly(A)⁺, 183–185
 - poly(T) probe for, 183–185
 - procedure for, 181–183
 - phenol for purification of, 588
 - structure of
 - effect of stringency on, 174f
 - factors that affect, 175–176, 175t
 - stringency of, 174–178
 - tips for maintaining, 586
 - ultraviolet spectrophotometry
 - determination of, 140–145
- Nucleic acid hybridization, 283, 501
 - factors influencing, kinetics and duplex stability, 284
 - formamide, 290
 - G+C content, 288
 - ionic strength, 286–287
 - mismatching, 288–289
 - pH, 287
 - probe complexity, 289
 - probe concentration, 287–288
 - probe length, 287
 - temperature, 285–286
 - urea, 290–291
 - viscosity, 289–290
 - and Northern analysis, 292–296
 - posthybridization stringency washes, 295–296
 - prehybridization
 - defined, 292
 - filter preparation, 293
 - for long probes, 293–294
 - probe denaturation, 294–295
 - probe removal and rehybridization, 296–298, 297t
 - rationale for, 283–284
 - temperature of, 291–292
- Nucleic acid molecules
 - double-stranded, 174
 - hybridization of, 175
 - stringency of, 174–178
- Nucleic acid pellets, 586
- Nucleic acid probe(s)
 - application of, 262
 - classification of, 263
 - definition of, 261
- DNA probes, 272
 - double-stranded, 272
 - first-strand cDNA, 272
 - 5'-end-labeling of, 275
 - labeling of, 273–275
 - nick translation, 274
 - oligonucleotides, 273
 - PCR synthesis of, 274
 - random priming, 274
 - synthesis, 273–275
 - 3'-end-labeling of, 275
- heterogeneous, 263, 272

- Nucleic acid probe(s) (*Continued*)
 homogeneous, 263
 labeling of, 263–264
 chemiluminescence and, 268–272
 Cy3 and Cy5, 267–268
 DNA probes, 264f
 isotopic, 265–267
 non-isotopic, 267
 purification, 279–280
 RNA, 276
 characteristics of, 278
 description of, 276–277
 5' end-labeling, 279
 hybridization, 276
 synthesis, 278–279
 3' end-labeling, 279
 in vitro transcription, 276f, 277f, 278–279
 storage, 280–281
 Nucleic acid sequence based amplification, 443
 Nucleoside, defined, 3
 Nucleotide
 assembly of, 5–8
 comparative structure of, 4
 nitrogenous bases in, 7
 Nuclide safe handling, 266
 Nylon filters, for Northern analysis
 description of, 244–245
 ultraviolet cross linking RNA to, 257
 Nylon macroarrays, 493, 501
- O**
 O.C.T. compound, 90
 OD. *See* Optical density
 -OH. *See* Hydroxyl group
 Oligodeoxythymidylic acid, 369
 Oligo(dT), 125
 Oligo(dT)-cellulose column
 chromatography, 129–134
 Oligonucleotide primers, 371, 400
 Oligonucleotide probes, 273, 501
 Oligonucleotides
 melting temperature for, 292
 sequence, 408
 Operon
 lac, 15
 polycistronic, 15
- Optical density, 130, 228
 spectrophotometric measurements
 of, 142
 Optitran, 244
Oryza sativa, 11
- P**
³²P, 265t, 266, 267
³³P, 265–266, 265t
 PABP. *See* Poly (A)-binding protein;
 Polyadenylate-binding protein
 PAGE. *See* Polyacrylamide gel
 electrophoresis
 Passive capillary transfer, for Northern
 analysis, 246–247, 250–252
 PBGD. *See* Porphobilibogen deaminase
 PBS. *See* Phosphate-buffered saline
 PCR. *See* Polymerase chain reaction
 PCR-based assays
 optimized reactions, sensitivity
 of, 477
 real-time- and end-point, 496
 PCR-based transcription assays, 450
 PCR products, melting curve analysis,
 463–464
 Peptidylprolyl isomerase A, 418
 pH
 nucleic acid hybridization and, 287
 for phenol saturation, 604–605
 stringency affected by, 176
 Phenol, 59
 description of, 588
 oxidation, 603–604
 preparation of, 603–606
 redistilled, 603–604
 saturation, 604–606
 Phenol:chloroform: isoamyl alcohol,
 166–167
 Phenolic lysate, pH of, 68
 Phosphate buffer, 599
 preparation of, 184t
 Phosphate buffered saline, 89, 599
 Phosphodiester backbone, 6, 256
 PhosphorImager, 232, 312,
 319–320
 Photodocumentation
 camera tips, 237
 definition of, 233
 digital systems for

- description of, 223–228, 233
- digitization process, 224
- example of, 223, 223f
- image analysis software used with, 223, 224f
- electrophoretograms, 235–238
- film
 - development of, 236–237
 - high-speed, with ultraviolet light, 236
 - recovery of, 238
- filtration used in, 234–235, 235t
- image analysis systems used with
 - advantages of, 229–230
 - background correction using, 230
 - benefits of, 229t
 - budget for, 231–232
 - image formats, 228–229
 - misconceptions regarding, 230
 - optical density, 228
 - PhosphorImagers, 232
 - practical considerations for, 229–231
 - software used with, 223, 224f
 - terminology associated with, 224–228
 - workshops, 232
- limitations of, 230, 238–239
- optimization of, 235–238
- rationale for, 221–222
- safety during, 222–223
- sample visualization, 234
- traditional systems for
 - description of, 233, 233f
 - limitations of, 230, 238–239
- ultraviolet light used in
 - film used with, 236
 - safety concerns, 222–223
- X-ray films, 238–239
- Photolithography, 501
- Picture element, 224, 225f
- Pixels, 224, 225f
- Plants cells
 - types of RNA produced in, 109–110
 - ultrastructural features of, 105
- Plateau effect, 389, 452
- Poly(A)-binding protein, 123
- Poly(A) caveat, 123–126
- Polyacrylamide gels
 - electroblotting used with, 249
 - electrophoresis, 410
 - concentration of, 630–631
 - description, 628
 - procedure for, 637–639
 - RNA denaturation using, 186
- Polyadenylate-binding protein, 28
- Polyadenylation, 122–123
- Polyadenylation signal, 28
- Poly(A)⁺ mRNA
 - nucleic acid normalization, 183–185
 - posthybridization washes, 185
 - purification of
 - advantages and disadvantages of, 125t
 - magnetic bead for, 127–129
 - oligo(dT)-cellulose column chromatography, 129–134
 - rapid, non-column, 135–136
 - selection of, 126–127, 199
 - slot blot normalization of, 182
 - yield of, 181
- Poly(A) tail, 122–123
 - of mRNA, 26–28
- Polylinker region, 438
- Polymerase chain reaction, 122, 385
 - adjuncts, effect of, 412
 - alternative transcription start sites, 434–435
 - amplification
 - enzyme blends, 431
 - gene-specific primers, 477
 - ligase chain reaction, 443
 - linear RNA amplification, 441–442
 - nucleic acid sequence based amplification, 443
 - secondary competitive, 478
 - strand displacement amplification, 442
 - analysis of, 420–421
 - annealing temperature, 406
 - asymmetric reaction, 388
 - blunt-end PCR products, 438–439
 - cDNA, amplification of, 436–438
 - classical methods, 386
 - competitive, 472–479
 - control reaction formats, 467–471
 - control reactions, 472

Polymerase chain reaction (*Continued*)

cycles, components of
 denaturation, 388–389
 primer annealing, 390
 primer extension, 390
 description, 576, 580
 detection methodology, 477
 DNA probe labeling and, 274
 enzymes, stability of, 413
 exogenous transcript, 465–467
 first-strand cDNA synthesis, 435–436
 genomic DNA/cDNA sequences,
 enzymatic amplification of, 449
 genomic-length, synthesis of, 429
 heteroduplexes, 476
 image analysis, 485–486
 internal controls, 416–419, 465
 ligation reactions, 439
 long-range, 430
 mechanics of, 387
 micropipette tips, 398
 micropipettors, 408
 negative control reactions, 471–472
 non-real-time competitive, 479–485
 optimization procedures, 408–411
 cycling profile, 416
 locked nucleic acids, 415–416
 negative controls, 413
 non-hot-start format, 413–415
 positive controls, 413
 thermostable polymerases, 411–413
 plateau effect, 452
 prevention
 aerosol-resistant tips, 397–398
 laboratory design, 395–396
 procedural methods, 396–397
 uracil-N-glycosylase, 398
 primer design. *See* Primers, PCR
 product inventory, 392
 quality control technique in, 152–153
 quantitative approaches, 451–453
 real-time, 421, 453–464
 LightCycler 2.0, 454
 LUX primers, 460–461
 melting curve analysis, 462–464
 Molecular beacon, 458–460
 Scorpions, 461–462
 SYBR green assay, 457
 TaqMan assay, 458

threshold cycle, 454
 reverse northern analysis, 424
 reverse transcriptase
 approach of, 391–395
 downstream gene-specific primer, 400
 process, 391–392
 quality control points, 421–423
rTth and *Taq* activities, 394
 second-strand cDNA synthesis, 394
 sensitivity of, 486–488
 Southern analysis, 407
 spurious product formation, 414
 TA cloning ligation reaction, 439–440
 techniques
 5' RACE, 424–425
 5' RLM Race, 426
 long-range, concept of, 429–431
 nested, 428–429
 single cell, 431–432
 splinkerette, 433
 3' RACE, 426–428
 TOPO cloning method, 440–441
 touchdown, 416–417
 transcription controls, 419–420
 transcription template, 469, 470
 transcript quantification, 473
 Polysomal mRNA, RNA isolation from,
 98–99
 Polysome fraction, 15
 Polysomes, 15
 Poly(T) probe, synthesis of, 183–185
 Polytron disruption, in homogenization
 method, 84–86
 Polyvinylidene difluoride, 211, 245
 Porphobilinogen deaminase, 419
 Portable PCR workstation, 471
 Posttranscriptional gene silencing, 542
 Potassium acetate, RNA isolation using,
 74–75
 PPIA. *See* Peptidylprolyl isomerase A
 Precipitation
 for nucleic acid, 50t
 rate of, 49
 Preflashing film, 309
 Prehybridization
 defined, 292
 filter preparation, 293
 for long probes, 293–294
 Pribnow box, 12

Primer-dimer formation, 462

Primers

analyzer, 404

biotinylated, 268–269

downstream

differential display PCR, 529–530

nomenclature, 399

oligonucleotide, 371, 400

PCR, 353, 449

annealing of, 390

basic rules, 402–405

design of, 398–402, 449

multiplex primer design, 407–408

nomenclature, 399

on-line resources, 407

qualities of, 387

recognition sequences of, 469

T_m considerations, 406–407

Probe(s)

application of, 262

classification of, 263

complexity, 289

concentration of, 287–288, 288t

definition of, 261

denaturation of, 294–295

DNA, 272

double-stranded, 272

3'-end-labeling of, 275

5'-end-labeling of, 275

first-strand cDNA, 272

labeling of, 273–275

nick translation, 274

oligonucleotides, 273

PCR of, 274

random priming, 274

synthesis, 273–275

heterogeneous, 263, 272

homogeneous, 263

length of, 287

nuclease protection assay

selection of, 327–329

single-stranded, 327

purification, 279–280

rehybridization, 296–298

removal, 296, 297t

method for, 297–298

RNA, 276

characteristics of, 278

description of, 276–277

5' end-labeling, 279

hybridization, 276

synthesis, 278–279

3' end-labeling, 279

in vitro transcription, 276f, 277f,
278–279

storage, 280–281

Prokaryotic gene, 15

Prokaryotic mRNA, 30–31

Prokaryotic RNA

isolation of, 75–76

Promoter

CAAT box, 13

defined, 11

eukaryotic, 13

in eukaryotic cells, 13

in silico analysis, 12

TATA box, 13

Pronase, 168

Proteinase K, 168

Protein biochips, 500

Protein chips, 495

Protein–protein interactions, 492

Proteins, isoelectric point, 495

Protein translation machinery, 341

Proteome

defined, 563

description of, 566–571

Proteomics

defined, 563

description of, 566–571

Psoralen biotin, 269

PTGS. *See* posttranscriptional gene silencing

Purification

of poly (A)⁺ mRNA, 125t

of polysome-engaged mRNA, 96–97

PVDF. *See* polyvinylidene difluoride

Pyrimidine, 3

Q

Quality control

northern analysis, 153

to support RT-PCR, 152–153

ultraviolet shadowing, 150–152

ultraviolet spectrophotometry for,
140–145

in vitro translation, 153

Quantifying transcript abundance, 451

Quinones, 603

- R**
- RACE. *See* Rapid amplification of cDNA ends
- Radiation Safety Office, 266
- Radioactive half-life, 266
- Radiolysis, defined, 229
- Random priming
DNA probe labeling and, 274
- Rapid amplification of cDNA ends, 425
- Rare earth screens, 307
- RdRP. *See* RNA-dependent RNA polymerase
- Real-time PCR, 421, 453–464
LightCycler 2.0, 454
LUX primers, 460–461
measurement, 455
melting curve analysis, 462–464, 463–464
microarrays and, 506
Molecular Beacon, 458–460
Scorpions, 461–462
standard curve, 456
SYBR green assay, 457
TaqMan assay, 458, 459
threshold cycle, 454
- Recordkeeping, 591–592
- Redistilled phenol, 603–604
- Reference gene, 122, 418
- Reference RNA, 510
- Relative abundance, 576
- Reverse genetics, 539–540
- Reverse northern analysis
advantage of, 259
convenience of, 259
description of, 258, 259f
- Reverse transcriptase-polymerase chain reaction
advantages/disadvantages, 474–475
amplification plot, 455
cDNA, 374
cRNA amplification, 467
DNA and RNA, 494
downstream gene-specific primer, 400
gene expression, 494
mRNA, 394
northern analysis *vs.* nuclease protection assay, 344
nuclease protection assay, 499
nuclease protection assay *vs.* nuclear run-off assay, 344
quality control points, 421–423, 422f
RNA molecules, 385
single cell, transcriptional activity, 432
standard curve, 456
- Ribonucleases, 362
A, 159
activity, 361
contamination, 45
description of, 155–156
electrophoresis contamination by, 213
elimination of activity of, 156–157
endogenous activity of, 156
extrinsic activity of, 156
H, 159, 373
incubation, for removing RNA from DNA sample, 613–614
intrinsic activity of, 156–157
latent, 157–158, 158f
pancreatic, 155
reagents and equipment for
CsCl, 167–168
CsTFA, 168
DEPC, 162–164
8-hydroxyquinoline, 167
GTC, 165–166
guanidine hydrochloride, 165
H₂O₂, 164
NaOH, 165
N-laurylsarcosine, 166
phenol:chloroform: isoamyl alcohol, 166–167
preparation of, 161–162
proteinase K, 168
RNA_{later}, 169
SDS, 165, 166
sterile water for irrigation, 164
- RNasin, 159–160
- T₁, 159
- types of inhibitors
nonspecific, 160–161
specific. *See* Specific inhibitors, of ribonucleases
- VDR. *See* Vanadyl ribonucleoside complexes
- Ribonucleoprotein particles, 109
- Riboprobes, 279
- Ribosomal RNA, 8, 148

- amount of, 181
- 18S, 198, 199
- 5S, 198
- molecular weight of, 194–204
- 28S, 198, 199
- RNA
 - adapter, 427
 - affinity matrices in, 73–74
 - for blot analysis, 88
 - capillary transfer of, setup for, 247f
 - competitor transcripts, 474
 - denaturation of
 - description of, 185–187
 - formaldehyde, 186–191
 - formamide, 187
 - glyoxal/dimethyl sulfoxide, 191–194
 - methylmercuric bromide, 187
 - selection of, 186
 - urea, 187, 189–191
 - deproteinization, 48
 - electrophoresis of, 192–194
 - electrophoretic profile of plant, 107f
 - excision of introns in, 17
 - expedient isolation of, 46
 - glyoxalated, 192
 - isolation of. *See* Lysis buffer
 - formulations
 - isopycnic centrifugation of, 64f
 - kit technology in, 71–74
 - membrane solubilization of, 47–48
 - molecular weight standards for. *See* Molecular weight standards, for electrophoresis
 - on non-denaturing gels, 194
 - nucleic acid concentration for, 48–49
 - nucleocytoplasmic transport, 46
 - purification of, 47–49
 - rationale for, 45–46
 - sensitivity index, 580
 - short hairpin RNA
 - comparative anatomy, 547f
 - defined, 542
 - mammalian cell delivery of, 546–547
 - mechanism of, 546–548
 - in vitro* production for, 547f
 - in vivo* transcription of, 548
 - vs.* siRNA, 544t
 - short interfering RNA
 - comparative anatomy of, 547f
 - defined, 542
 - design of, 550–552
 - mammalian cell delivery, 548–549
 - in vitro* synthesis of, 546f
 - vs.* shRNA, 544t
 - steps involved in, 543–544
 - defined, 542
 - design of, 550–552
 - mammalian cell delivery, 548–549
 - in vitro* synthesis of, 546f
 - vs.* shRNA, 544t
 - silica technology in, 72–73
 - simultaneous isolation of, 68–71
 - species, 345
 - storage of, 78–79
 - study of, 1–2
 - synthesis, 348, 492
 - gene expression, 500
 - total inhibition of nuclease activity, 48
 - transcripts, 343, 346, 360
 - types of, 8–9
 - types of, in plant cells, 109–110
 - ultraviolet spectrophotometry
 - determination of, 140–145
- RNA-dependent RNA polymerases, 22, 359
- RNA hairpins, 186, 391
- RNAi. *See* RNA interference
- RNA-induced silencing complex, 543, 544
 - defined, 542
- RNA interference
 - application, 557
 - defined, 539–540
 - description of, 186
 - mechanism of, 542–550
 - miRNA, 549–550
 - nomenclature associated with, 541–542
 - rationale, 539–541
 - short hairpin RNA
 - comparative anatomy, 547f
 - defined, 542
 - mammalian cell delivery of, 546–547
 - mechanism of, 546–548
 - in vitro* production for, 547f
 - in vivo* transcription of, 548
 - vs.* siRNA, 544t
 - short interfering RNA
 - comparative anatomy of, 547f
 - defined, 542
 - design of, 550–552
 - mammalian cell delivery, 548–549
 - in vitro* synthesis of, 546f
 - vs.* shRNA, 544t
 - steps involved in, 543–544

- RNA interference (*Continued*)
 validation
 Northern analysis, 556
 RT-PCR approaches, 556
 Western analysis, 556–557
- RNA isolation
 from bark, 112–113
 from fruit, 113–115
 guanidinium buffers, 57–58
 kit for, 71–74
 from leaf, 110–112
 from plant, 105–109
 potassium acetate for, 74–75
 prokaryotic, 75–76
 strategies for plant tissues, 115–116
 from tissue
 “clean up” methods, 100–101
 fixed tissue, 91–92
 fresh tissue, 89–90
 frozen tissue, 90–91
 LiCl-urea method, 92–95
 lipid-enriched, 95–96
 polysomal mRNA, 98–99
 troubleshooting of, 101–102
 troubleshooting from plant tissue,
 116–117
 from yeast, 76–78
- RNA-later, 99–100, 169
- RNA molecules, *in vitro* transcription of,
 465, 468
- RNA pellets, 586
- RNA polymerases, 12, 18–22, 345, 346
 activities of, 359–360
 eukaryotic enzymes in, 18
 genes, transcription of, 342
 nucleic acid synthesis of, 19
- RNA probes, 276
 characteristics of, 278
 description of, 276–277
 3' end-labeling of, 279
 5' end-labeling of, 279
 hybridization, 276
 synthesis, 278–279
in vitro transcription, 276f, 277f,
 278–279
- RNA purification, 361
- RNase. *See* Ribonucleases
- RNase protection assay, 323
 characteristics of, 325
 protocol for, 335–337
 RNase A and RNase T1, comparison
 of, 326t
 RNA transcript quantification using,
 324f
 S1 nuclease assay *vs.*, 326t
- RNA sequences, 345
- RNasin, 378
RNasin, 159–160
- RPA. *See* RNase protection assay
- RpoT gene, 109
- rRNA. *See* Ribosomal RNA
- RT-PCR. *See* Reverse transcriptase-
 polymerase chain reaction
- RUBISCO, 110
- Running buffer, 213
- S**
- ³⁵S, 265, 265t, 266
- Safelight, 306
- Safety
 during electrophoresis, 212–213
 during photodocumentation, 222–223
- Saline sodium citrate, 599
- Salt
 electrophoresis affected by, 216f
 stringency affected by, 176
- Salting-out, for high-molecular-weight
 DNA isolation, 621–623
- Sarkosyl, 166
- scRNA. *See* Small cytoplasmic RNA
- SDA. *See* Strand displacement amplification
- SDS. *See* Sodium dodecyl sulfate
- Second-strand cDNA synthesis reaction
 cloning, 379
 PCR, 394
- Sensitivity assay
 highest, 451
 lower, 450, 457
- Sensitivity index
 RNA, 580
- Sephadex G-25, 280
- Sephadex G-50, 280
- Sequence-specific restriction enzyme, 440
- 7-methylguanosine (m⁷G), 24
- Short hairpin RNA
 comparative anatomy, 547f
 defined, 542
 mammalian cell delivery of, 546–547

- mechanism of, 546–548
- in vitro* production for, 547f
- in vivo* transcription of, 548
- vs.* siRNA, 544t
- Short interfering RNA
 - comparative anatomy of, 547f
 - defined, 542
 - design of, 550–552
 - mammalian cell delivery, 548–549
 - in vitro* synthesis of, 546f
 - vs.* shRNA, 544t
- Silanizing, 617
- Silica-based purification products, 379
- Silica technology
 - to purify RNA, 72–73
- Silver staining, 209–210
- Single-cell PCR, 432
- Single-codon amino acids, 401
- Single nucleotide polymorphisms, 401, 460, 492, 500
- Single-stranded DNA, 141
- SI units, 664
- 6-carboxyfluorescein, 458
- 6-FAM. *See* 6-carboxyfluorescein
- Small cytoplasmic RNA, 17
- Small nuclear ribonucleoproteins, 17, 25
 - overview of, 35–36
- SNP. *See* Single nucleotide polymorphisms
- snRNP. *See* Small nuclear ribonucleoproteins
- S1 nuclease assay
 - characteristics of, 325
 - disadvantages of, 325
 - protocol for, 332–334
 - RNA transcript quantification using, 323f
 - RPA *vs.*, 326t
- Sodium chloride, 598
- Sodium dodecyl sulfate, 46, 165, 166, 411
 - description of, 588
 - preparation of, 599
- Sodium hydroxide, 165, 598
- Solution hybridization, 575, 579
- Southern analysis, 261, 262
- Southern blotting, 180, 242
- Specific inhibitors, of ribonucleases
 - RNasin, 159–160
 - VDR. *See* Vanadyl ribonucleoside complexes
- Splinkerette PCR, 433
- Spun-column chromatography, 280
- SSH. *See* Suppression subtractive hybridization
- Steady-state RNA, 346
- Sterile water for irrigation, 164
- Stock solution, 597–601
 - mass to moles, conversion of, 593–595
- Stop bath solutions, 304
- Storage of purified RNA
 - long-term, 78–79
 - short-term, 78–79
- Strand displacement amplification, 442
- Streptavidin, 269, 270, 313
- Streptomyces avidinii*, 269, 313
- Streptomyces griseus*, 168
- Stringency
 - definition of, 174
 - effect of formamide on, 177
 - effect of pH on, 176
 - effect of salt on, 176
 - effect of temperature on, 177
 - effect of urea on, 177–178
 - importance of controlling, 174–178
 - posthybridization washes, 295–296
- Subtraction
 - troubleshooting, 524–526
 - Suppliers, 655–662
- Suppression subtractive hybridization, 496
 - adaptor-ligated tester cDNAs, 518, 520f
 - description, 515–516
 - driver in, 516
 - efficiency of, 516
 - hybridizations in, 520–521
 - plasmid preparation, 521, 523
 - replica plated tool used with, 523, 523f
 - results of, 522f
 - schematic diagram of, 519f, 521f, 522f
 - tester in, 516
 - tester subpopulations used in, 518
- SYBR Gold, 207
- SYBR Green, 286, 457
 - description of, 207, 208t, 209f
 - disposal of, 607–609
- SYBR Safe, 208–209

- T**
- TA cloning, 438
 description, 437
 ligation reaction, 439–440
 schematic diagram of, 438f
- TAP. *See* Tobacco acid pyrophosphatase
- TaqMan assay, 458, 459
- Taq polymerase, 273, 390, 393, 429
- Target, 261
- TATA boxes, 13, 434
- TCA precipitation, step and filtration, 352
- T-cell receptor genes, 493
- TE buffer, 600
- TE-9 buffer, 600
- Technology overload, 232
- Temperature
 hybridization, 291–292
 melting
 description of, 177
 formamide effects on, 175
 nucleic acid hybridization and, 285–286
 stringency affected by, 177
- Template strand, 14
- Terminal deoxynucleotidyl transferase, 425
- Terminal transferase, 264, 264f
- Tester, 516
- Tetracycline, 600
- Tetramethylethylenediamine, 638
- TFPA-PEG 3 -Biotin, photo-reactive, 269
- Thermal cycler, for PCR, 587
- Thermostable polymerase, 412
- Thermodynamic stability, 291
- Thermus aquaticus*, 390
- Thermus thermophilus*, 373
- Thin layer chromatography, 122
- 3-aminophthalate, 316, 317f
- 3'-end-labeling
 of DNA probes, 275
 of RNA probes, 279
- Thymine, 3, 3f
- TIFF files, 229
- Tissue
 homogenization methods
 description of, 84
 Dounce homogenizer, 86–87
 motorized devices for, 84
 polytron disruption, 84–86
 RNA isolation from
 “clean up” methods, 100–101
 fixed tissue, 91–92
 fresh tissue, 89–90
 frozen tissue, 90–91
 LiCl-urea method, 92–95
 lipid-enriched, 95–96
 polysomal mRNA, 98–99
 troubleshooting of, 101–102
- Tissue culture
 advantages of, 82–83
 tissue samples *vs.*, 83
- TLC. *See* Thin layer chromatography
- Tobacco acid pyrophosphatase, 426
- Tonoplast, 106
- TOPO cloning, 381, 440
- Total cellular RNA, size distribution of, 59f
- Toxin α -amanitin, 18
- Trademarks, 667–669
- Trailer sequence, 26
- Transcript abundance, interpolation of, 468
- Transcription
 animal models for, 11
 of mRNA, 10–11
 plant models for, 11
 products of, 18–22
in vitro, 14
- Transcription assays
 mechanism of, 343
 nuclear run-off assay, 343–345
 nuclear run-off *vs.* nuclear run-on assay, 346–347
 sensitivity index, 580
 steady-state RNA, 346–347
- Transcription start site, 2, 12, 383, 424, 427, 434
- Transcriptome, 565–566
 defined, 563
- Transcriptomics, 492, 565–566
 defined, 564
- Transcript profiling, 492
- Transfer RNA, 8, 147, 186
de novo protein synthesis in, 8
- Translation, 11
 affect in mRNA sequence and structure, 31–32
 comparison of, 20
- Tris, 600–601
- Tris-acetate-EDTA, 600

Tris-borate-EDTA, 600
 Tris/SDS buffer, 601
Tritirachium album, 168
 tRNA. *See* Transfer RNA
 Trypsinization protocol for
 anchorage-dependent cells,
 619–620
 TSS. *See* Transcription start site
 TurboBlotter, 247–248, 248f, 252–254
 Turgor pressure, in plant, 106
 Tuschl's rules, 551
 28S rRNA, 198, 199
 2'-deoxyribose sugar, 3
 2,5- diphenyloxazole, 308

U

Ultrafree filtration units, 280
 Ultraviolet light
 crosslinking of RNA using, 256–257
 photodocumentation uses of, 222–223
 Ultraviolet spectrophotometry for
 absorption ratios, 140–145
 UNG. *See* Uracil-N-glycosylase
 UNG enzyme, 398
 Universal reference RNA, 509
 Upstream, 7
 Uracil, 3f, 5
 Uracil-N-glycosylase, 398
 Urea, 290–291
 effect of, on stringency, 177–178
 RNA denaturation using, 187, 189–191
 URR. *See* Universal reference RNA
 UTR. *See* 5' untranslated region
 UV irradiation, 397
 UV light. *See* Ultraviolet light

V

Vaccinia virus enzyme topoisomerase I,
 440
 Vacuum blotting, 248–249

Vanadyl ribonucleoside complexes
 contraindications, 159
 development of, 158
 ethylenediaminetetraacetic acid removal
 of, 159
 preparation of, 159
 structure of, 159
 synthesis of, 169–170
 VDR. *See* Vanadyl ribonucleoside
 complexes
 Very low abundance, 22, 122, 181
 Viscosity, nucleic acid hybridization and,
 289–290

W

Water, 601
 Watson-Crick base-pairing, 7
 Web sites, 292, 571, 582
 Western analysis, 263

X

XAR film, 305
 Xenobiotic-induced recessive mutations,
 497
 X-ray film, 453
 autoradiography uses of, 305–306,
 305t
 developing scheme for, 303f
 exposure time for, 306–307
 safelight for, 306
 photodocumentation uses of,
 238–239

Y

YAC. *See* Yeast artificial chromosomes
 Yeast
 culture medium, 601
 RNA isolation from, 76–78
 Yeast artificial chromosomes,
 564–565