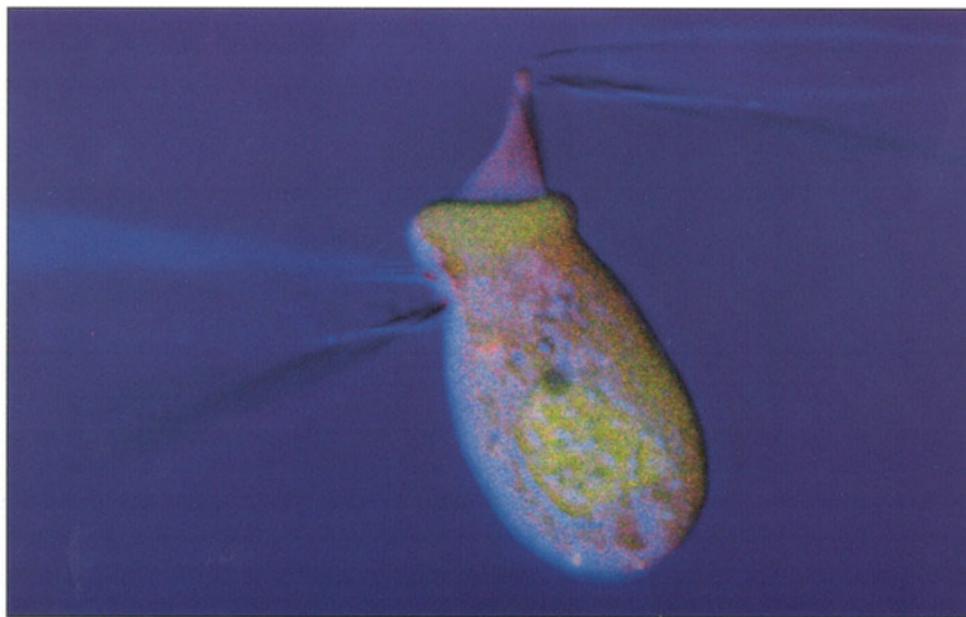


Colloquium on
Auditory Neuroscience:
Development, Transduction,
and Integration

National Academy of Sciences



Edited by A. James Hudspeth and Masakazu Konishi

COLLOQUIUM ON Auditory Neuroscience: Development, Transduction, and Integration

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C.

NATIONAL ACADEMY OF SCIENCES

In 1991, the National Academy of Sciences (NAS) inaugurated a series of scientific colloquia, several of which are held each year under the auspices of the NAS Council's Committee on Scientific Programs. Each colloquium addresses a scientific topic of broad and topical interest, cutting across two or more traditional disciplines. Typically two days long, colloquia are international in scope and bring together leading scientists in the field. Papers from colloquia are published in the *Proceedings of the National Academy of Sciences* (PNAS).

National Academy of Sciences Colloquia Bound Reprints Available



In 1991, the National Academy of Sciences (NAS) inaugurated a series of scientific colloquia, several of which are held each year under the auspices of the NAS Council Committee on Scientific Programs. These colloquia address scientific topics of broad and topical interest that cut across two or more traditional disciplines. Typically two days long, these colloquia are international in scope and bring together leading scientists in the field.

Papers presented at these colloquia are published in the *Proceedings of the National Academy of Sciences* (PNAS) and are available online (www.pnas.org). Because they have generated much interest, these papers are now available in the form of collected bound reprints, which may be ordered through the National Academy Press.

Currently available are:

Carbon Dioxide and Climate Change (\$11)

Held November 13–15, 1995 (Irvine, CA)

Computational Biomolecular Science (\$16)

Held September 12–13, 1997 (Irvine, CA)

Earthquake Prediction (\$16)

Held February 10–11, 1995 (Irvine, CA)

Elliptic Curves and Modular Forms (\$7)

Held March 15–17, 1996 (Washington, DC)

Genetic Engineering of Viruses and Viral Vectors (\$21)

Held June 9–11, 1996 (Irvine, CA)

Genetics and the Origin of Species (\$8)

Held January 31–February 1, 1997 (Irvine, CA)

Geology, Mineralogy, and Human Welfare (\$11) New!

Held November 8–9, 1998 (Irvine, CA)

Neurobiology of Pain (\$8) New!

Held December 11–13, 1998 (Irvine, CA)

Neuroimaging of Human Brain Function (\$17)

Held May 29–31, 1997 (Irvine, CA)

Plants and Population: Is There Time? (\$8) New!

Held December 5–6, 1998 (Irvine, CA)

Protecting Our Food Supply: The Value of Plant Genome Initiatives (\$13)

Held May 29–31, 1997 (Irvine, CA)

Proteolytic Processing and Physiological Regulation (\$11) New!

Held February 20–21, 1999 (Irvine, CA)

Science, Technology, and the Economy (\$12)

Held November 20–22, 1995 (Irvine, CA)

The Age of the Universe, Dark Matter, and Structure Formation (\$13)

Held March 21–23, 1997 (Irvine, CA)

Papers from future colloquia will be available for purchase after they appear in PNAS.

Shipping and Handling Charges:

In the U.S. and Canada please add \$4.50 for the first reprint ordered and \$0.95 for each additional reprint.

Ordering Information:

Telephone orders will be accepted only when charged to VISA, MasterCard, or American Express accounts.

To order, call toll-free 1–800–624–6242 or order online at www.nap.edu and receive a 20% discount.



Graeme Yates
1944–2000

Graeme Yates, an Australian auditory physiologist and internationally recognized expert on cochlear function, died on October 13 after a courageous struggle with a long illness. Graeme had been invited to speak in the NAS Colloquium on *Auditory Neuroscience: Development, Transduction, and Integration*, but was too sick to attend. His penetrating insights into both mammalian and non-mammalian cochlear function were sorely missed. Graeme was at the height of his power as a scientist and was planning new studies of cochlear mechanics and neurophysiology until only days before his death. Over 25 years of publications attest to his major contributions to our current understanding of cochlear mechanics and afferent neural output, to the analysis of otoacoustic emissions, and to descriptions of two-tone interactions and adaptation. Graeme made multifaceted contributions to the many research teams of which he was part, investigating the hearing of mammals, reptiles, and birds, designing hardware, writing software, and providing fundamental theoretical contributions to data interpretation. All those who have had the privilege of working with Graeme Yates cherish their memories of this civilized man and passionate scientist. His death is a tragedy for auditory physiology, and his loss will be felt acutely by his friends and colleagues.

PNAS

Proceedings of the National Academy of Sciences of the United States of America

Contents

COLLOQUIUM

Papers from the National Academy of Sciences Colloquium on Auditory Neuroscience: Development, Transduction, and Integration

INTRODUCTION

Auditory neuroscience: Development, transduction, and integration 11690

A.J.Hudspeth and Masakazu Konishi

COLLOQUIUM PAPERS

Notch signaling in the development of the inner ear: Lessons from *Drosophila* 11692

Mark Eddison, Isabelle Le Roux, and Julian Lewis

Molecular genetics of pattern formation in the inner ear: Do compartment boundaries play a role? 11700

John V.Brigande, Amy E.Kiernan, Xiaoying Gao, Laurie E.Iten, and Donna M.Fekete

Patterning of the mammalian cochlea 11707

Raquel Cantos, Laura K.Cole, Dario Acampora, Antonio Simeone, and Doris K.Wu

Cellular studies of auditory hair cell regeneration in birds 11714

Jennifer S.Stone and Edwin W.Rubel

Hair cell recovery in mitotically blocked cultures of the bullfrog sacculle 11722

Richard A.Baird, Miriam D.Burton, David S.Fashena, and Rebecca A.Naeger

Two mechanisms for transducer adaptation in vertebrate hair cells 11730

Jeffrey R.Holt and David P.Corey

Cochlear mechanisms from a phylogenetic viewpoint 11736

Geoffrey A.Manley

Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: 11744**Comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chin-chilla**

Mario A.Ruggero, S.Shyamla Narayan, Andrei N.Temchin, and Alberto Recio

The spatial and temporal representation of a tone on the guinea pig basilar membrane 11751

K.E.Nilsen and I.J.Russell

Molecular mechanisms of sound amplification in the mammalian cochlea 11759

Jonathan F.Ashmore, Gwénaëlle S.G.Géléoc, and Lene Harbott

Putting ion channels to work: Mechano-electrical transduction, adaptation, and amplification by hair cells 11765

A.J.Hudspeth, Y.Cho, A.D.Mehta, and P.Martin

Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus 11773

Donata Oertel, Ramazan Bal, Stephanie M.Gardner, Philip H.Smith, and Philip X.Joris

Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus 11780

Jane J.Yu and Eric D.Young

Cellular mechanisms for resolving phase ambiguity in the owl's inferior colliculus 11787

José Luis Peña and Masakazu Konishi

Subdivisions of auditory cortex and processing streams in primates 11793

Jon H.Kaas and Troy A.Hackett

Mechanisms and streams for processing of "what" and "where" in auditory cortex 11800

Josef P.Rauschecker and Biao Tian

The corticofugal system for hearing: Recent progress 11807

Nobuo Suga, Enquan Gao, Yunfeng Zhang, Xiaofeng Ma, and John F.Olsen

Traces of learning in the auditory localization pathway 11815

Eric I.Knudsen, Weimin Zheng, and William M.DeBello

Plasticity in the neural coding of auditory space in the mammalian brain 11821

Andrew J.King, Carl H.Parsons, and David R.Moore

Spatial processing in the auditory cortex of the macaque monkey 11829

Gregg H.Recanzone

Song selectivity and sensorimotor signals in vocal learning and production 11836

Michele M.Solis, Michael S.Brainard, Neal A.Hessler, and Allison J.Doupe

On cortical coding of vocal communication sounds in primates 11843

Xiaoqin Wang

A new view of language acquisition 11850

Patricia K.Kuhl

Introduction

AUDITORY NEUROSCIENCE: DEVELOPMENT, TRANSDUCTION, AND INTEGRATION

A.J.Hudspeth*[†] and Masakazu Konishi[‡]

*Howard Hughes Medical Institute and Laboratory of Sensory Neuroscience, The Rockefeller University, 1230 York Avenue, New York, NY 10021–6399; and [‡]Division of Biology 216–76, California Institute of Technology, Pasadena, CA 91125

Hearing underlies our ability to locate sound sources in the environment, our appreciation of music, and our ability to communicate. Participants in the National Academy of Sciences colloquium on Auditory Neuroscience: Development, Transduction, and Integration presented research results bearing on four key issues in auditory research. How does the complex inner ear develop? How does the cochlea transduce sounds into electrical signals? How does the brain's ability to compute the location of a sound source develop? How does the forebrain analyze complex sounds, particularly species-specific communications? This article provides an introduction to the papers stemming from the meeting.

We live in a world of sounds. Although we often attend to these signals only subconsciously, hearing constantly informs us about our surroundings: people entering and leaving the room, equipment beginning and ending its tasks, announcements and alarms alerting us to change and danger. From plainsong to Smashing Pumpkins, audition underlies one of life's chief pleasures, the enjoyment of music. Most importantly, our communication with one another rests primarily on our ability to interpret the complex sonic signals that constitute speech. The study of hearing is therefore motivated not only by intellectual curiosity but also by an appreciation of the sense's importance in daily life and an interest in restoring hearing in those deprived of its virtues.

The National Academy of Sciences colloquium on Auditory Neuroscience: Development, Transduction, and Integration, held on May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA, reviewed recent progress in auditory research. Rather than attempting a comprehensive overview of the field, the colloquium's organizers sought to elicit contemporary answers to four questions. How is the ear formed? How does it transduce sounds into electrical signals? How does the brainstem develop its capacity to compute the spatial location of sound sources? How do the upper reaches of the auditory pathway analyze complex sounds? The balance of this article establishes the motivation for each of these queries and provides a précis of our current understanding.

Development of the Inner Ear

The ear's elaborate structure—justifiably called the labyrinth—forms from a simple slab of epithelial cells, the otic placode of the embryo. Developmental biologists have begun to elucidate the steps in this process. Cellular expression of a battery of morphogenetic proteins partitions the aural primordium into precursors for six receptor organs (1). In a series of origami-like steps, the otic cyst then folds into the three toroidal semicircular canals, the ellipsoidal utricle and saccule, and the snail-like cochlea. The constituent cells meanwhile begin to adopt several fates. Cells in the sensory patch of each receptor organ hone their identities by molecular competition with one another, yielding in the mature ear a crystalline array of hair cells separated by supporting cells. Incipient hair cells then erect their elaborate hair bundles by complex manipulations of the cytoskeleton (2). Supporting cells simultaneously differentiate into several distinct types whose functions remain obscure. After neuroblasts have left the sensory epithelium, the daughters of their cell divisions coalesce into ganglia adjacent to the labyrinth. The resultant neurons innervate hair cells and extend axons along the eighth cranial nerve into the brain, where they transmit information to cells of the cochlear and vestibular nuclei.

Because hair cells in the human cochlea are not mitotically replaced, their number declines throughout life as a result of genetic abnormalities, ear infections, loud sounds, ototoxic drugs, and aging. As a consequence, about one-tenth of the population in industrialized countries suffers from significant hearing loss. Research on the development of hair cells is accordingly motivated in part by the expectation that an understanding of the factors involved in creating hair cells will suggest a means of regenerating them. There are several reasons to hope for success in this endeavor. First, it is clear that supporting cells can serve as hair-cell precursors: in fishes and amphibians, hair cells are formed throughout life by this means. Next, functional hair cells have been shown to regenerate in avian cochleas after destruction of the original receptors with loud sounds or ototoxic drugs. Finally, several growth factors have already proven effective in promoting the mitosis of hair-cell precursors in the mammalian utricle. If new hair cells can be created in the human cochlea, their potential connection to the nerve fibers surviving nearby offers an excellent opportunity for the restoration of hearing.

Transduction of Stimuli in the Inner Ear

Not only can we hear sounds of frequencies from 20 Hz to 20 kHz, but a trained musician can discriminate frequencies with a precision of $\approx 0.1\%$. An important topic of research for over a century therefore has been the mechanism by which stimulus frequency is represented along the basilar membrane. Our understanding of this process rests on three fundamental insights. First, as adduced by Helmholtz (3), each increment of the approximately 30-mm-long basilar membrane is tuned to a particular frequency by such mechanical properties as its mass and tension. Next, as demonstrated by von Békésy (4), sound energy flows through the fluids of the cochlea, producing a traveling wave along the basilar membrane. Finally, as hypothesized by Gold (5), the cochlea contains an active element that

[†]To whom reprint requests should be addressed. E-mail: hudspaj@rockvax.rockefeller.edu.

This paper is the introduction to the following papers, which were presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held on May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

amplifies mechanical inputs and allows resonant responses despite the damping effects of viscosity.

The nature of the amplifier that mediates cochlear sensitivity and frequency discrimination is a topic of lively debate. Mechanical amplification originated over 350 million years ago, for it occurs in amphibians and in all four ramifications of the amniote vertebrates (6). At least in nonmammalian tetrapods, amplification seems to result from active movements of the mechanoreceptive hair bundles (7). Mammals have evolved a distinctive amplificatory appurtenance, the outer hair cell. Electrical stimulation of this cell causes it to elongate or contract, a movement thought to effect amplification by pumping energy into the basilar membrane's oscillation (8). It remains unclear whether this electromotile mechanism has supplanted hair-bundle motility as the amplificatory mechanism in mammals, whether the two processes coexist, or whether electromotility serves another purpose altogether.

Recent studies of cochlear mechanics have used laser interferometry to provide details of the basilar membrane's elaborate motion. As a result of saturation in the cochlear amplifier, the structure's responsiveness is highly nonlinear. The peak sensitivity occurs for threshold sounds, which elicit movements of less than ± 1 nm; amplification is negligible for loud sounds (9). The waveform of oscillation suggests that, even near the threshold, dozens of outer hair cells contribute to the amplification of a pure sinusoidal input. Interferometric measurements also imply that each increment of the basilar membrane does not simply oscillate up and down, but rather that the inner and outer portions of the membrane move in opposite directions as the overlying tectorial membrane resonates independently.

Processing of Sound in the Brainstem

The auditory system is built for speed. Hair cells transduce stimuli in microseconds, a striking contrast to the tens to hundreds of milliseconds required by photoreceptors and olfactory neurons. Axons in the auditory nerve can fire action potentials at rates approaching 1,000 per second. Specialized glutamate receptors speed synaptic processing along the auditory pathways; the lavish use of K^+ channels lowers neuronal time constants and shortens the climb to threshold (10). In keeping with this intense signaling activity, histochemical staining reveals that the auditory system has the highest metabolic rate in the brain.

The rapidity and temporal precision of auditory processing underlie one of the fundamental functions of the auditory brainstem, the localization of sound sources in space. Like many other animals, we often detect a novel environmental feature by hearing it, then turn our eyes or head for closer inspection. Although quite routine, this procedure involves remarkable neural computations. Interaural time difference, the delay in the arrival of a sound at the ear farther from its source relative to that at the nearer ear, is a key clue to a sound source's position. But even a sound coming directly from one side reaches the near ear only 600 μ s earlier than the far one, an interval comparable to the duration of a single action potential. Our finest discrimination of a source's position involves measurement of interaural time delay with a precision of less than 20 μ s—a seemingly impossible feat that we reflexively perform dozens of times a day.

Our ability to localize sound sources is not confined to the horizontal dimension; we can also situate an aural target along the vertical axis. Here the corrugated surface of the external ear is of prime importance, for the efficiency with which the pinna captures sounds originating at different elevations depends on their frequencies. The dorsal cochlear nucleus appears to be the neural computer charged with inferring sound-source elevation from the resultant spectral clues.

For the brain's sound-localization apparatus to direct eye and head movements, it is essential that an exact correspondence exist between the sensory representations of sound sources and of visual objects. This interaction in fact occurs when a map of auditory space, created by neurons in the inferior colliculus, projects to the optic tectum or superior colliculus of the midbrain to form a bimodal, visual-auditory map (11). In both owls and ferrets, developmental studies indicate that the visual map regulates the auditory one: after derangement of the correspondence by respectively offsetting visual images with prisms or deflecting the eyes by surgery, the auditory map shifts so as to regain its congruence with the visual map. Studies of this elegant form of neural plasticity have now pinpointed the site where the shift occurs, which in owls lies in the external nucleus of the inferior colliculus.

Analysis of Complex Sounds by the Forebrain

The most important role of hearing in our daily lives is the perception of speech. Audition likewise serves many other animals in the analysis of signals from conspecifics: the alarm calls of numerous gregarious species, the territorial and mate-attracting songs of birds, and the extensive lexicon of primates. The processing of communication signals is very difficult, as attested in the instance of human speech by the fact that computers have achieved a limited degree of success only after 50 years' effort! The neuronal substrate for analysis of complex sounds, including those associated with conspecific communication, is beginning to emerge from contemporary investigations.

Although the auditory cerebral cortex of primates has been known for decades to occupy the dorsal surface of the temporal lobe, the complexity of the region has been appreciated only recently. The auditory cortex now is known to have at least 15 subdivisions, each with distinct patterns of anterograde and retrograde projection. Although neurons in the core region of the auditory cortex are responsive to pure-tone stimuli, those in the belt of surrounding cortical areas are better activated by more complex sounds, including species-specific vocalizations.

Among the most discriminating auditory areas studied to date are the telencephalic nuclei of the song system in songbirds. Neurons here respond only to species-specific song and distinguish between song syllables played in different orders. Studies on this topic, as well as on the cortical analysis of species-specific calls in primates (12), are especially exciting because they seem likely to shed light on the mechanism of our most profound auditory ability, the interpretation of speech.

We thank Dr. J.Halpern and Mr. K.Fulton for initiating the colloquium, Mr. E.Patte for administrative assistance, and the National Academy of Sciences for financial support. We are especially grateful to Ms. Beth Dougherty of The Rockefeller University for organizing the meeting and to Ms. M.Gray-Kadar for supervising the excellent meeting facilities of the Arnold and Mabel Beckman Center.

1. Fekete, D.M. (1999) *Trends Neurosci.* **22**, 263–269.
2. Kollmar, R. (1999) *Curr. Opin. Neurobiol.* **9**, 394–398.
3. Helmholtz, H. (1954) in *On the Sensations of Tone* (Dover, New York), pp. 139–148.
4. von Békésy, G. (1960) in *Experiments in Hearing* (McGraw-Hill, New York), pp. 403–534.
5. Gold, T. (1948) *Proc. R.Soc. London Ser. B* **135**, 492–498.
6. Manley, G.A. & Köppl, C. (1998) *Curr. Opin. Neurobiol.* **8**, 468–474.
7. Hudspeth, A.J. (1997) *Curr. Opin. Neurobiol.* **7**, 480–486.
8. Nobili, R., Mammano, F. & Ashmore, J. (1998) *Trends Neurosci.* **21**, 159–167.
9. Ruggero, M.A. (1992) *Curr. Opin. Neurobiol.* **2**, 449–456.
10. Trussell, L.O. (1999) *Annu. Rev. Physiol.* **61**, 477–496.
11. Knudsen, E.I. & Brainard, M.S. (1995) *Annu. Rev. Neurosci.* **18**, 19–43.
12. Doupe, A.J. & Kuhl, P.K. (1999) *Annu. Rev. Neurosci.* **22**, 567–631.

Colloquium

NOTCH SIGNALING IN THE DEVELOPMENT OF THE INNER EAR: LESSONS FROM *DROSOPHILA*

Mark Eddison*[†], Isabelle Le Roux^{†‡}, and Julian Lewis*[§]

*Vertebrate Development Laboratory and [†]Developmental Genetics Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, United Kingdom

The sensory patches in the ear of a vertebrate can be compared with the mechanosensory bristles of a fly. This comparison has led to the discovery that lateral inhibition mediated by the Notch cell-cell signaling pathway, first characterized in *Drosophila* and crucial for bristle development, also has a key role in controlling the pattern of sensory hair cells and supporting cells in the ear. We review the arguments for considering the sensory patches of the vertebrate ear and bristles of the insect to be homologous structures, evolved from a common ancestral mechanosensory organ, and we examine more closely the role of Notch signaling in each system. Using viral vectors to misexpress components of the Notch pathway in the chick ear, we show that a simple lateral-inhibition model based on feedback regulation of the Notch ligand Delta is inadequate for the ear just as it is for the fly bristle. The Notch ligand Serrate1, expressed in supporting cells in the ear, is regulated by lateral induction, not lateral inhibition; commitment to become a hair cell is not simply controlled by levels of expression of the Notch ligands Delta1, Serrate1, and Serrate2 in the neighbors of the nascent hair cell; and at least one factor, Numb, capable of blocking reception of lateral inhibition is concentrated in hair cells. These findings reinforce the parallels between the vertebrate ear and the fly bristle and show how study of the insect system can help us understand the vertebrate.

Almost all animals, from cnidarians (1) to mammals, have mechanosensory organs for touch and detection of vibrations and other disturbances of the air or water in which they live. This sensory capability, it seems, is as important and as universal as sensitivity to light, suggesting that the apparatus of mechanosensation, like that of photoreception (2), may have a very ancient evolutionary origin. Mechanosensory organs such as the ear may be elaborate and highly specialized according to the animal's way of life, but at their core they must always have a set of mechanosensory transducer cells to perform the fundamental task. At the level of these cells and their immediate companions one may hope to find conserved features reflecting evolution from a common prototype: homologous cell types, homologous developmental processes, and homologous molecular mechanisms.

In the vertebrate ear, the core structures are the sensory patches, consisting of hair cells (the transducers), supporting cells (which form the epithelial framework in which hair cells are held), and the adjacent cochleovestibular sensory neurons (which synapse with the hair cells). How are these cell types, especially the hair cells and supporting cells, generated in the correct pattern and proportions? Parallels with *Drosophila* provide a route toward an answer: the sensory patches in the vertebrate ear have a counterpart in the sensory bristles of the fly, suggesting that homologous mechanisms may operate (3, 4). This approach has revealed that lateral inhibition mediated by the Notch signaling pathway (5), a key mechanism for controlling cell diversification in fly sense organs, has a similarly crucial role in the vertebrate inner ear.

In the first part of this paper we review the published evidence on Notch signaling in the ear and see how it fits with a simple model (6) that has been proposed to explain the patterning of the ear's sensory patches (7). This model does not, however, correspond accurately to the way in which Notch signaling governs development of sensory bristles in the fly (8–10). Are the vertebrate ear and the insect bristle not so closely homologous after all, or is the simple model proposed for the ear misleading? We review the arguments for homology, and in the second half of the paper we present experimental evidence to test the role of Notch signaling in the ear. We show that Notch signaling in the ear is more complex than was originally suspected: at least three Notch ligands are at work; they are regulated in contrary and complementary ways; and, contrary to previous suggestions, the pattern of cell determination cannot simply be explained in terms of lateral inhibition rules governing Notch-ligand expression. These corrections to previous ideas, far from undermining the arguments for homology with *Drosophila* sensilla, make the parallels seem even closer than before.

The Hair Cells, Supporting Cells, and Neurons of a Sensory Patch Have a Common Origin in the Otic Placode. The inner ear derives from the otic placode, a thickening of the epidermis adjacent to the hindbrain in the early embryo (11–13). This placode gives rise both to the inner ear epithelium, with its sensory patches consisting of hair cells and supporting cells, and to the sensory neurons that innervate these patches. The neuronal lineage becomes segregated from the sensory epithelial lineage at an early stage: as the otic placode invaginates and forms first a cup and then a vesicle, neuroblasts become singled out within the otic epithelium and begin to delaminate from its anteroventral portion (4). These cells will divide a variable number of times before differentiating as cochleovestibular neurons. In the epithelium, the future sensory patches become identifiable by their expression of markers such as *Serrate1* (*Ser1*) and *BMP4* (4, 14), but differentiation of hair cells and supporting cells does not begin until 3–5 days later (in birds and mammals).

Although the lineage relationship between the neuroblasts and the hair and supporting cells has not been rigorously determined, it seems that the anteroventral region of the early otocyst is a source for all three cell types, while a more posteromedial region gives rise only to hair cells and supporting cells (4). It has, however, been shown that hair cells and supporting cells have the same ancestry and frequently arise as

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: DI1, Delta1; Ser1, Serrate1; Ser2, Serrate2; Su(H), Suppressor of Hairless; *En*, embryonic day *n*; SMC, sensory mother cell; GFP, green fluorescent protein; HCA, hair cell antigen; Lfng, Lunatic fringe.

[†]M.E. and I.L.R. contributed equally to this work.

[§]To whom reprint requests should be addressed. E-mail: j.lewis@icrf.icnet.uk.

pairs of sister cells (15–17). Given that these two cell types have a common origin and lie intermingled in the otic epithelium, sharing a common environment, what causes them to be different? How is the correct mixture of hair cells and supporting cells generated?

The Alternating Mosaic of Hair Cells and Supporting Cells Suggests a Simple Lateral Inhibition Model for Genesis of Spatial Pattern. The ratio of hair cells to supporting cells in a mature sensory patch is variable from one region to another. In the auditory epithelium (the basilar papilla) of the 12-day chicken embryo, for example, it ranges from 1:1.7 at the distal end to 1:3.9 at the inferior-proximal end (18). Despite this variation, the distribution of hair cells among supporting cells seems almost everywhere to obey the same simple rule: with few exceptions, every cell that lies in contact with a hair cell is a supporting cell, and any cell that escapes all contact with hair cells is itself a hair cell. This immediately suggests a pattern-generating mechanism based on lateral inhibition (Fig. 1) (3, 19). According to this hypothesis, the cells in the undifferentiated sensory patch have a choice of two fates (hair cell or supporting cell), and the choice is governed by interactions between nearest neighbors. The default or primary fate, in the absence of interaction, is to become a hair cell, but in the interacting system each nascent hair cell delivers an inhibitory signal to the cells in contact with it, deterring these from becoming hair cells, too, and at the same time preventing them from producing inhibitory signals that would act back on the nascent hair cell. Neighbors thus compete to be hair cells, and the losers in the competition, the cells surrounding hair cells, become supporting cells.

Lateral inhibition of this sort is well documented in *Drosophila* and *Caenorhabditis elegans*, and studies in these species have revealed the molecular mechanism that transmits the inhibitory signal (5, 20). The receptor on the cell receiving inhibition is the transmembrane protein Notch, and the ligand on the adjacent cell that delivers inhibition is the transmembrane protein Delta. Loss-of-function mutations in the Delta-Notch signaling pathway abolish lateral inhibition and allow an excessive proportion of the population to adopt the primary fate.

On this basis, a simple formal model for pattern generation can be proposed (21–23): an increase of Delta in one cell causes increased activation of Notch in its neighbor, and Notch activation in the neighbor down-regulates Delta expression in that same cell, as well as inhibiting commitment to the primary fate (Fig. 1). The feedback control of Delta expression has the effect of amplifying contrasts between adjacent cells: a sheet of initially similar cells, with only small random differences between them, will spontaneously develop into a mosaic of alternate cell types of just the type seen in the sensory patches of the ear (6).

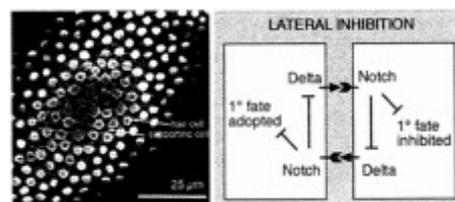


Fig. 1. (Left) An optical section (confocal image) in the plane of the chick basilar papilla, showing the mosaic of hair cells (stained with HCA antibody, white blobs) and supporting cells (outlined by their cortical actin, stained with fluorescent phalloidin). (Reproduced from ref. 4.) (Right) Lateral inhibition with feedback: a simple model of Delta-Notch signaling that can generate spatial pattern. Only two cells are shown, interacting to adopt two different fates, but the same principles apply to a field of many cells, which by interacting in this way can in theory give rise to the type of pattern seen at Left.

Lateral Inhibition in the Sensory Patches of the Ear Is Mediated by the Notch Signaling Pathway. In the vertebrate central nervous system, or at least in the neural plate (24, 25) and retina (26), the production of neurons from neuroepithelial progenitor cells seems to be controlled in just the way that the simple lateral-inhibition-with-feedback model proposes. All of the progenitors express the Notch homolog *Notch1*, while nascent neurons appear within this population as scattered cells expressing the Delta homolog *Delta1* (*Dll*) at higher levels than their neighbors. When *Dll* is artificially overexpressed, so that all cells strongly deliver and receive the inhibitory signal, the cells all are prevented from differentiating as neurons. Conversely, when all of the cells are forced to express a dominant-negative form of *Dll*, blocking Delta-Notch signaling, they all escape inhibition and differentiate as neurons prematurely. Moreover, expression of the endogenous *Dll* gene is regulated in accordance with the simple feedback model. All of the requirements for pattern generation by lateral inhibition with feedback, and all of the predictions of the model, seem to be satisfied.

To see whether the same is true in the ear, the first step is to examine the normal gene expression patterns in the developing sensory patches (4, 27, 28). *Notch1* is expressed, as in the central nervous system, throughout the population of progenitor cells (and in fact even more widely in the otic epithelium). *Dll* is expressed in a scattered subset of the population, and these *Dll*-expressing cells can be identified as the nascent hair cells (4, 27). These observations are backed up by similar circumstantial evidence for Notch signaling during hair-cell regeneration in birds (63).

Direct functional evidence comes from zebrafish and mice. Thus in zebrafish with a mutation in the *deltaA* gene—one of four zebrafish *Delta* homologs, all normally expressed in nascent hair cells—hair cells are produced in increased numbers, as though inhibition has been reduced (29). Likewise, there is some overproduction of hair cells in mice with a knockout mutation in a gene coding for another Notch ligand, *Serrate2* (also known as *Jagged2*, but henceforward in this paper *Serrate2* to match chick and *Drosophila* terminology) (28). *Serrate2* (*Ser2*) is a homolog of the *Drosophila Serrate* gene, closely related to *Delta* and likewise coding for a protein capable of activating Notch; in the sensory patches of the ear it is expressed with *Dll* in the nascent hair cells (although more persistently).

The phenotypes of the *deltaA* and *Ser2* mutations are relatively mild but consistent with the model. The zebrafish *mind bomb* mutant provides more spectacular evidence. This mutant takes its name from its neurogenic phenotype: in the central nervous system, it grossly overproduces neurons at the expense of neural progenitor cells, in a manner that is diagnostic of a failure of lateral inhibition. The excess of neurons goes with an up-regulation of *Delta* expression: the cells are deaf to the inhibitory signal that Delta normally delivers. In the ear, the result of this failure of Notch signaling is dramatic: each of the prospective sensory patches differentiates in its entirety into a uniform mass of hair cells, with no supporting cells (7). The hair cells are not only produced in great excess, as much as 30-fold, but also disappear within a day or two after their first appearance (30).

The *mind bomb* phenotype tells us that Notch signaling is required (i) to prevent the cells in a developing sensory patch from all differentiating alike as hair cells, and (ii) to delay production of hair cells until the proper time. It shows that lateral inhibition mediated by the Notch pathway is essential, but it falls short of proving that lateral inhibition with feedback is the

mechanism that generates the normal orderly mixture of hair cells and supporting cells. Indeed, there are several reasons to be cautious in adopting this simple model mechanism, seductive as it may seem. Goodyear and Richardson (18) have shown, for example, that the orderliness in the basilar papilla of the chick depends to a large extent on cell rearrangements occurring after cell differentiation. While *Dll* and *Ser2* are concentrated in nascent hair cells, there is yet another Notch ligand, *Ser1*, that is strongly expressed in supporting cells (4, 27, 31). Most significantly, in the sensory bristles of *Drosophila*, Delta expression is not regulated in the way the model postulates, even though lateral inhibition via Notch is as essential as in the ear (8, 9). Is the *Drosophila* bristle a misleading paradigm, or is the simple theoretical model wrong? To answer these questions, we first review the parallels between insect bristles and the sensory patches of the vertebrate ear.

The Sensory Patches of the Vertebrate Ear Resemble the Sensilla of a Fly in Function and Developmental History. Each insect bristle is a miniature sense organ, or sensillum, consisting of a set of four cells: a neuron, a neural sheath cell, a bristle socket cell, and a bristle shaft cell. These cells normally all derive, along with a migratory glial cell (32), from a single sensory mother cell, or SMC. Lateral inhibition mediated by Notch operates repeatedly in development of the bristle, first to single out the SMC within a cluster of competent cells in the epidermis, and then at each subsequent cell division to drive the diversification of the progeny of the SMC to form the differentiated cells of the bristle (8, 10, 33, 34).

The insect bristle resembles the sensory patch in the vertebrate ear in several respects. First of all, the mechanosensory function is essentially the same, with the same kinetics of response and adaptation (35).

Second, the cell types correspond, in part at least: neuron with neuron, bristle shaft cell with hair cell, bristle socket cell with ear supporting cell. Bundles of actin filaments form the shaft of the bristle shaft cell, just as bundles of actin filaments form the stereocilia of the hair cell (36). Both these cell types have a well-defined planar polarity, essential for directional sensitivity.

Third, the developmental anatomy is similar. Just as the component cells of the bristle have a common origin in the epidermis, so do the neurons, hair cells, and supporting cells of the ear have a common origin in the otic placode ectoderm. Just as the first division of the SMC gives rise to a neuronal precursor, which delaminates from the epidermis, and an epithelial precursor, which stays behind to generate the socket and shaft cells, so also the first step in differentiation of the sensory patch is production of neuroblasts, which delaminate from the otic epithelium, and sensory epithelial precursors, which stay behind to form hair cells and supporting cells. In both systems, the final differentiated cell types are the products of a series of dichotomous cell-fate choices, and this series of choices is similar in the fly sensillum and the vertebrate ear.

Lastly, and most crucially, the molecular mechanisms underlying these choices are similar.

Corresponding Events in Development of Insect Sensillum and Vertebrate Inner Ear Are Controlled by Homologous Systems of Genes. In fly sensilla, basic helix-loop-helix (bHLH) transcription factors of the Achaete/Scute family and the related Atonal family have a key role, both in initiating the program of sensory development (the “proneural” function) and as differentiation factors for the final cell types (22). Examples of the latter include the products of the *asense* and *cousin of atonal (cato)* genes, which drive differentiation both in bristles and chordotonal organs (a closely related type of sensillum) (37). In the ear, although bHLH genes serving the early proneural function have not been identified, it has been shown that the *atonal* homolog *Math1* is expressed selectively in hair cells and drives hair-cell differentiation: hair cells fail to develop in a *Math1* knockout mouse (38), and nonsensory cells in the cochlea differentiate into hair cells when transfected with *Math1* (39). Upstream from the proneural genes in the fly, controlling their domains of expression, lie transcription factors of the Iroquois family; these, too, have vertebrate homologs that are expressed at early stages in the future sensory epithelium of the ear (40).

Components of the Notch signaling pathway lie downstream from the proneural genes and play a central and recurrent role in the subsequent development of the fly sensillum. In each of the series of cell-fate decisions in the sensillum lineage, from the singling out of the SMC onward, lateral inhibition mediated by Notch signaling is required (8, 33, 34).

The Notch signaling pathway has a similarly central and recurrent role in the development of the sensory patches of the vertebrate ear. *Notch1* is expressed throughout the otic placode at the very beginning of ear development and thereafter throughout the developing sensory epithelium, persisting into adult life in the supporting cells (4, 16). *Dll* is expressed at each of the sites where cell-fate choices are being made (4). During the segregation of neural and epithelial sublineages of the ear, *Dll* RNA is seen in scattered cells in the neurogenic region of otic epithelium, apparently the nascent neuroblasts. Subsequently it is expressed in the ganglion formed by the neuroblasts, and, as we have seen, in the sensory-patch epithelium as hair cells are being generated and becoming different from supporting cells. The block of Notch signaling in the *mind bomb* mutant not only causes overproduction of hair cells, but also neurons (7).

Finally, the molecular data reveal additional similarities between the differentiated cell types. In particular, the Pax gene *D-pax2* is expressed specifically in the bristle shaft cell and required for its correct differentiation (41); the vertebrate homolog *Pax2* is expressed in the early otic epithelium and then selectively in hair cells and also is required for their correct differentiation. In the mouse *Pax2* knockout, no cochlea forms (42), and in a zebrafish *pax2.1* mutant, hair-cell differentiation is abnormal (29).

The parallels between insect bristles and ear sensory patches that we have summarized above add up to a persuasive argument that these mechanosensory organs are indeed homologous—that they resemble one another because they have evolved from a common ancestral prototype. There are, of course, also important differences, as one might expect after 800 million years of divergent evolution. The fly bristle is ensheathed in a semirigid cuticle, making the mechanics of stimulus delivery quite different from that in the ear. In the ear, the hair cells are the transducers and synapse with the neurons; in the bristle, there is no synapse and the dendrite of the neuron acts as the transducer. In the development of the ear, there is no step corresponding to the singling out of the SMCs from the epidermis: instead, the future sensory patch behaves like a uniform mass of contiguous SMCs. The numbers of cell divisions elapsing between one cell-fate choice and the next are different in the two systems, and the cell lineage patterns are not identical. Almost all of these developmental differences, however, correspond to plausible evolutionary variations and have parallels in differences between the various, but unmistakably related, types of sensilla in the fly (see ref. 4 for details).

The Pattern of Cell Types in the Insect Bristle Is Not Controlled Through Regulation of Delta Expression. Notch signaling in the developing bristle shows three important departures from the simple lateral-inhibition model sketched in Fig. 1. First, recent studies have shown that Delta is not the only Notch ligand at work. Serrate is also present and must be mutated along with Delta to give the most extreme Notch-pathway loss-of-function phenotype (43). Second, Delta is not regulated in the way the simple model

postulates, either at the mRNA or at the protein level. Delta expression remains high in the winning and losing cells during many of the lateral-inhibition interactions that decide cell fate (8, 9, 34), implying that some other factor(s) must create a difference by modulating the efficacy of Delta or the ability to respond to it. Third, factors acting in this way to bias Notch signaling have been identified. In particular, the intracellular protein Numb is distributed asymmetrically between the daughters and granddaughters of the SMC and makes the cells that contain it deaf to lateral inhibition (10, 44).

We now turn to our experimental observations on Notch signaling in the sensory patches of the chick ear.

Materials and Methods

Viral Constructs. The RCAS-D11 and RCAS-D11^{dn} replication-competent retroviral constructs were as described (26). RCAS-X-Su(H)^{dn} was a gift from J.-C. Izpisua-Belmonte (Salk Institute, La Jolla, CA) and contains a form of the *Xenopus* Suppressor-of-Hairless [Su(H)] cDNA with its DNA binding mutated, as described (45). The viruses were used at a titer of 5×10^7 – 10^9 cfu/ml.

The pseudotype virus will be described in detail elsewhere (I.L.R., unpublished work). Briefly, plasmids based on the LZRSpBMN-Z plasmid (46) were prepared by inserting cDNA for the product of interest (D11 or D11^{dn}) linked to an internal ribosome entry site followed by DNA coding for green fluorescent protein (GFP); this composite coding sequence was placed under the control of a 253-bp upstream enhancer sequence from Rous sarcoma virus, within LZRSpBMN-Z. Pseudotype virus then was generated by transiently cotransfecting 293gp packaging cells (Qiagen, Chatsworth, CA) with this construct plus a plasmid coding for vesicular stomatitis virus (VSV)-G protein. The resulting pseudotype virus particles contain RNA coding for D11+GFP or D11^{dn}+GFP, with Gag, Pol, and VSV-G proteins provided by the packaging cells. Virus released into the supernatant was concentrated by ultracentrifugation to a final titer of 5×10^8 – 10^9 cfu/ml.

Embryos and Chicken Injection. Chick embryos were windowed at stage 13 [embryonic day 2 (E2)], and $\approx 0.5 \mu\text{l}$ of virus solution (with $0.8 \mu\text{g}/\mu\text{l}$ polybrene, 3% methyl cellulose, and a trace of fast green dye) was injected into the lumen of the otic cup. Embryos were fixed between E6 and E9.

In Situ Hybridization and Immunohistochemistry. *In situ* hybridization was performed on 15- μm cryosections of fixed embryos as described (26), by using FastRed (Boehringer) for detection by fluorescence. For immunostaining, the cryosections were incubated overnight at 4°C in blocking solution (PBS with 3% BSA, 10% FCS, 0.1% Triton) containing the primary antibody. For Numb, sections were taken through an additional methanol fixation before the primary antibody was added. Chick anti-Numb antibody (47) was a gift from Y. Wakamatsu, Tohoku University, Japan. Ser1 was detected with a polyclonal antiserum as described (4). D11 and D11^{dn} were detected with a rabbit polyclonal antiserum directed against amino acids 325–462 of chick D11 (26). In the specimens shown here, this antiserum was used at a concentration sufficient to detect the high levels of exogenous D11 or D11^{dn}, but not endogenous D11. Hair cells were detected by using the hair-cell antigen (HCA) antibody (48). GFP was detected with a rabbit polyclonal antiserum (gift from D. Shima, Imperial Cancer Research Fund, London). Secondary antibodies were labeled with Alexa488 or Alexa594 (Molecular Probes). Images were collected by confocal microscopy. A total of ≈ 180 virus-injected embryos were serially sectioned and analyzed. Results are based on 47 embryos in which we saw informative patches of infection, i.e., patches that overlapped or touched sensory patches in the ear epithelium.

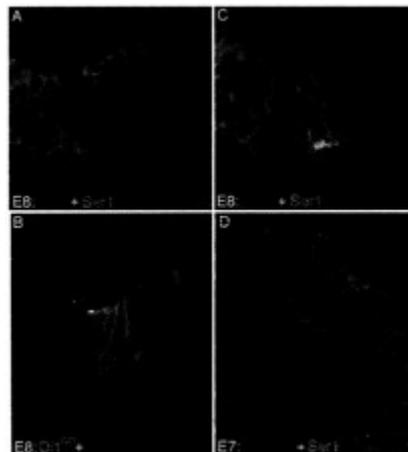


Fig. 2. Blocking Notch signaling causes down-regulation of Ser1. (A–C) Infection with RCAS-D11^{dn} virus, blocking signaling via Notch. Three consecutive sections through the utricle of an 8-day (E8) embryo are shown. The expression of D11^{dn} is shown by *in situ* hybridization with a *D11* probe (red) (A and C) and by antibody staining (green) (B). Ser1 expression is shown by antibody staining (green) (A and C); the distribution of hair cells is shown with the HCA antibody (detecting hair bundles, red) (B). Ser1 expression is lost at sites of infection. Presence of hair cells proves that these sites lie within a sensory patch, where normally Ser1 is expressed. (D) Infection with RCAS-XSu(H)^{dn}. Infection was detected by antibody against viral gag protein (red) and Ser1 expression by Ser1 antibody (green). Again, the two stains are mutually exclusive: cells infected within the sensory patch down-regulate Ser1.

Results

Ser1 Is Regulated by Lateral Induction. During the development of sensory patches in the chick ear, the Notch ligand Ser1 is expressed in a pattern quite different from that of D11 and Ser2. From a very early stage, it marks all of the cells in the prospective sensory patches, and although it is eventually down-regulated in hair cells as they differentiate, it remains high in all of the supporting cells (31). Because these are in contact with one another, this finding suggests that the expression of Ser1 is not regulated by lateral inhibition, which would tend to make levels different in neighboring cells. In the *Drosophila* wing margin, it has been found, however, that Notch activation also can regulate Notch ligand expression in a contrary way, by lateral induction, so that neighboring cells stimulate one another to express ligand strongly (49, 50). The pattern we see in the ear suggests that, while D11 and Ser2 may be regulated by lateral inhibition, Ser1 may be regulated by lateral induction.

To test this, it is necessary to interfere with Notch activation levels and see how Ser1 expression is altered. This can be achieved in the chick by infecting the cells with a viral vector that drives expression of a molecule that blocks Notch signaling. We have used two constructs, both based on the replication-competent RCAS virus (see *Materials and Methods*). One, RCAS-D11^{dn}, contains the dominant-negative truncated form of *D11*, which makes cells that contain it deaf to Notch signaling (26). The other blocks Notch signaling by interfering with Su(H), a transcription factor through which Notch acts (5): the viral construct, RCAS-X-Su(H)^{dn}, contains a mutated vertebrate homolog of Su(H) with a defective DNA binding domain; it thus also serves to block Notch signaling (45).

Embryos were injected with virus at 2 days and fixed 4–7 days

later. Patches of infection within sensory regions of the otic epithelium were analyzed for Ser1 expression (Fig. 2). Ser1 expression was lost or clearly reduced in 8 of 10 patches of infection with RCAS-D11^{dn}, and in 8 of 10 patches of infection with RCAS-X-Su(H)^{dn}. A control for nonspecific effects of viral infection was provided by a parallel series of experiments (see below) using another virus, RCAS-D11, containing the fulllength form of *D11*. In the majority of these cases (20 of 25 patches) no down-regulation of Ser1 was seen. Taken together, these data indicate that Ser1 is indeed positively regulated by Notch activity.

Effects of D11^{dn} Expression on Hair Cell Differentiation Are Difficult to Decipher. It might be expected that hair cells should be seen in excess in regions where Notch signaling was blocked with RCAS-D11^{dn} or RCAS-X-Su(H)^{dn}. We have examined hair cell production in our experiments with RCAS-D11^{dn}, in which embryos were infected at 2 days of incubation and fixed 4 or 6 days later. Altogether, we found 33 potentially informative infected patches. Strikingly, just over half (17/33) of the informative patches of infection directly abutted sensory patches, without any overlap (Fig. 3A and B): infection with the virus and sensory character appeared to be mutually exclusive. By comparison, in parallel experiments using the RCAS viral vector to misexpress full-length D11 (see below and Fig. 4), only 10% of informative infected patches abutted sensory patches in this way, as against 90% that overlapped with or were internal to sensory patches. At least two interpretations are possible. It could be that blockade of Notch signaling by RCAS-D11^{dn} converted prospective sensory patch cells to a nonsensory character. Alternatively, the loss of Notch activity may have caused premature hair-cell differentiation as in *mind bomb*, and as in *mind bomb* this may have lead to early death and disappearance of any infected cells that lay in a prospective sensory patch (7, 30). Where an early infection partially overlapped a prospective sensory patch, the loss of cells in the region of overlap would bring infected nonsensory cells into juxtaposition with uninfected sensory cells.

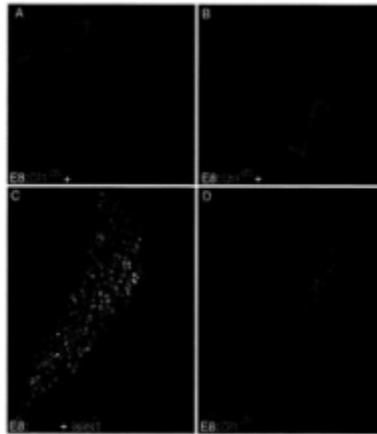


Fig. 3. Infections with RCAS-D11^{dn} virus. (A and B) Sections through utricle at E8; the D11^{dn}-expressing patches abut but do not overlap sensory patches. Two examples are shown. D11^{dn} expression is shown with an antibody against D11 (green), hair cells with HCA antibody (red). (C and D) D11^{dn}-expressing cells often end up as neurons in the cochleovestibular ganglion. (C) Section of basilar papilla and underlying ganglion at E6 stained by *in situ* hybridization for D11^{dn} (red) and with Islet1/2 antibody (green) to mark the nuclei of neurons in the cochleovestibular ganglion. (D) Adjacent section stained with D11 antibody, revealing the dendrites of infected neurons with D11^{dn} in their membranes.

Patches of infection with RCAS-D11^{dn} were seen inside sensory patches, or overlapping them, in 16 of 33 informative cases. Contrary to expectation, none of these patches showed an excessive density of hair cells (see Fig. 2B): 0.22 ± 0.07 hair cells were counted per μm length of sectioned epithelium in the infected regions, as compared with 0.23 ± 0.05 in the adjacent uninfected sensory tissue (mean \pm SD, $n=11$ in both cases; counts from representative sections showing patches $\geq 20 \mu\text{m}$ wide). Again, several interpretations are possible. Blocking Notch activity with RCAS-D11^{dn}, although it affects Ser1 expression, may fail to affect cell differentiation (different Notch family members with different sensitivities to D11^{dn} and different downstream actions could be involved, for example, or Ser1 expression and cell differentiation might have different thresholds of response to Notch activity). Alternatively, the cells may have been already irreversibly committed as hair cells or supporting cells before they became infected. Lastly, the cells may

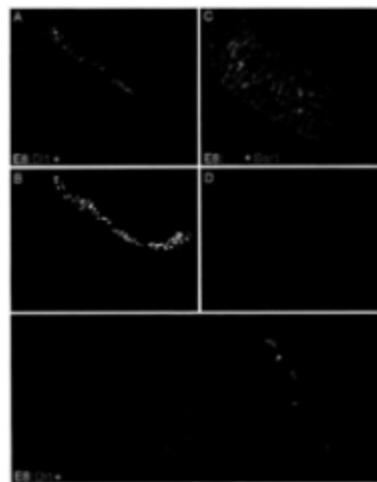


Fig. 4. Overexpression of D11 in a sensory patch does not inhibit hair-cell production. (A-D) Sections of utricle at E8, infected with RCAS-D11 virus. (A) Infected patch stained with an antibody against D11 (green) and with HCA antibody (red), (B) The same scene, showing the HCA signal only, to reveal the distribution of hair cells more clearly. (C) Adjacent section stained by *in situ* hybridization for D11 (red) and with Ser1 antibody. (D) Same scene as in C, with the red fluorescence hidden to show that D11, in contrast with D11^{dn}, does not cause down-regulation of Ser1. (E) Specimen infected with pseudotype (replication-defective) virus expressing D11+GFP, and stained with antibody against GFP (green) plus HCA antibody (red). Hair cells are produced normally even within clusters of contiguous cells all expressing D11.

not have been irreversibly committed, but may only recently have become infected by the replication-competent virus as it spread through the tissue and may not have had time yet to change their phenotype. More experiments will be needed to decide between these interpretations.

A further observation is more straightforward in conformity with the expected effects of blocking Notch activity. Infected cells expressing $D11^{dn}$ often ended up as neurons in the cochleovestibular ganglion (Fig. 3 C and D). Indeed, in 6 of a set of 6 embryos fixed at E6, infected cells were found only in the ganglion. This finding suggests that the infected cells were biased toward a neuronal fate during the early phase of ear development when neuroblasts delaminate from the otic epithelium. Virus-infected nerve fibers innervating a sensory patch of epithelium are a possible source of late infection for the epithelial cells.

Ectopic Expression of D11 Does Not Inhibit Hair-Cell Production. There are two ways in which a cell might escape lateral inhibition so as to become a hair cell: its neighbors might not deliver inhibitory signals, or it might be deaf to signals that it receives. In the sensory patches of the ear, all cells are in contact with neighbors expressing Ser1, suggesting that the nascent hair cells are deaf to at least this Notch ligand. They are not, however, normally exposed to D11, because D11 is not expressed by the supporting cells. To test whether forced expression of D11 throughout the cell population would block cell differentiation, as it does in the retina, we once again used viral vectors, this time containing the full-length *D11* cDNA. We found 29 RCAS-D11-infected patches that lay within sensory patches or overlapping them (Fig. 4 A-D), and three that lay directly abutting sensory patches. We counted the numbers of hair cells per unit length of sectioned epithelium in infected as compared with adjacent uninfected sensory epithelium. No significant difference was seen. The values were respectively 0.23 ± 0.08 and 0.24 ± 0.08 hair cells per μm (mean \pm SD, $n=18$ in both cases; counts from representative sections showing patches $\geq 30 \mu\text{m}$ wide).

The observations using RCAS-D11 to misexpress D11 were confirmed by using a pseudotype replication-defective virus for the same purpose. This virus gives smaller patches of infected cells, but has the advantage that they can be assumed to have all become infected at the same early time. Again, hair cells could be seen to develop normally even where all of the cells in the neighborhood expressed D11 (Fig. 4E).

The cells that become hair cells therefore do so regardless of whether their neighbors express D11. The simple lateral-inhibition model based on regulation of D11 expression cannot be the correct explanation of why some cells escape inhibition to become hair cells and others do not. Some other factor must operate, either interacting with Notch ligands in the neighbors of the nascent hair cell and blocking their ability to deliver an inhibitory signal, or interacting with the Notch pathway in the nascent hair cell and blocking its ability to respond.

Numb Protein May Make Nascent Hair Cells Deaf to Notch Signaling. Almost nothing is known about molecules that might interact with Notch ligands in cis to prevent them delivering a signal to an adjacent cell (see *Discussion*). Analogies with *Drosophila* suggest, however, at least three factors that might act in nascent hair cells to make them insensitive to signals received. First, the hair cells might down-regulate their expression of Notch1 itself— indeed, they are known to do so, although this may occur too late to control the cell-fate decision (16). Second, Delta protein in large quantities can make cells that contain it unresponsive to signals from neighbors (51), and hair cells contain D11. Lastly, there are proteins such as Numb that interact with Notch to block its activity (51). We have used immunohistochemistry to look for expression of a chick Numb homolog in the inner ear. As shown in Fig. 5, the chick Numb protein is localized to the basolateral membranes of the epithelial cells at stages before hair-cell differentiation. Once hair cells have differentiated, it is seen at high concentration in hair cells. These observations are preliminary. They clearly suggest, however, that Numb protein acts in nascent hair cells to make them immune to Notch signaling.

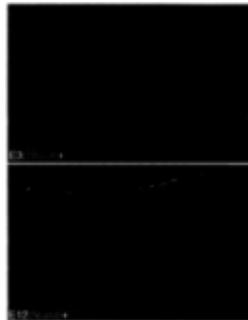


Fig. 5. C-Numb expression during sensory patch development. (A) Section of otic epithelium at E3, stained with Numb antibody (green) and counterstained with the nuclear dye Syto1 6 (red); note basal localization of Numb, apparently in all cells, including those undergoing mitosis close to the lumen. (B) Section of basilar papilla at E12, stained with Numb antibody (green) and HCA (red); the hair cells preferentially contain Numb, and it is no longer basally localized.

Discussion

The implications of our experiments are summarized diagrammatically in Fig. 6, showing the patterns of expression of the various Notch ligands in a newly differentiated sensory patch and their regulatory interactions. Supporting cells express Ser1; hair cells express D11 and Ser2. Ser1 expression is regulated positively by Notch activity; D11 and Ser2 are regulated negatively (7). Supporting cells contact one another, so that mutual lateral inductive signals, as well as signals from hair cells, keep them all in a state of high Notch activation, which maintains high expression of Ser1 and low expression of D11 and Ser2. Hair cells contain Numb and

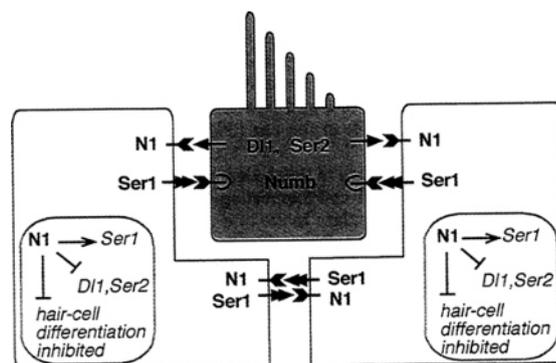


Fig. 6. Summary diagram of Notch-mediated interactions in a newly differentiated sensory patch in the chick ear. See *Discussion* for commentary.

down-regulate their Notch1 expression; thus despite their exposure to Ser1 from all sides, their level of Notch activation is low, their expression of Ser1 is low, and their expression of Dll1 and Ser2 is high. The cells that differentiate as hair cells do so because their level of Notch activation is low; those that differentiate as supporting cells do so because their level of Notch activation is high.

This system of regulatory interactions is robust and consistent with all of the observations. Our account is, however, based on some tacit assumptions and raises several questions that need to be discussed. The first concerns Ser1.

Signaling by Ser1 May Serve to Prevent Premature Hair-Cell Production. An attractive possibility is that Ser1, expressed throughout the prospective sensory patch from a very early stage, serves to prevent premature or ectopic hair-cell production by maintaining a high background level of Notch activation. This suggestion is supported by the recent finding that hair cells are overproduced in mouse cochlea explants when Ser1 (Jag1) production is inhibited with antisense oligonucleotides (52). Ser1 thus may be important in enforcing the long delay that occurs between neuroblast delamination and hair-cell differentiation in birds and mammals and in preventing the peripheral regions of growing sensory patches from differentiating prematurely in fish.

The supporting cells in sensory patches also express Lunatic fringe (Lfng) (4, 31), a homolog of the Fringe protein that has been shown in *Drosophila* to interact with Notch in cis and make Notch refractory by activation by Serrate (53, 54). This might be taken to suggest that Ser1 function normally is blocked by Lfng, in which case removal of Lfng should drastically alter the pattern of Notch activation and cell differentiation. However, ear development appears entirely normal in mice with a knockout mutation of the *Lfng* gene (55), and we see no abnormalities in ears overexpressing Lfng as a result of infection with RCAS-Lfng retrovirus (I.L.R., unpublished work). An effect of the *Lfng* knockout is seen only in mice that also have a mutation in *Ser2*: the *Ser2* single mutants show mild overproduction of hair cells, and this effect is partially suppressed in the double *Lfng; Ser2* mutants. Zhang *et al.* (54) interpret this in terms of an effect of Lfng as inhibitor of signaling by Dll1, although there is no precedent for such an effect. An alternative interpretation is that Lfng normally acts as a weak inhibitor of signaling by Ser1, so that loss of Lfng weakly potentiates the Notch-activating effect of Ser1; this might be expected to show up as a decrease of hair-cell production in the *Ser2* mutant background, where absence of one Notch ligand makes the functions of the others more critical.

More Than One Notch Homolog May Be Involved in Ear Development. A questionable assumption is that the various Notch ligands all act on the same receptor, Notch1. In fact, it has been reported that Notch2 and Notch3 are expressed along with Notch1 in the mouse ear (56, 57), and the different Notch family members might respond selectively to the different ligands and exert different downstream effects. Notch3, for example, may act as an antagonist of Notch1 (58). In the chick, however, only two Notch genes have been identified, and Notch1 is the only one that we have been able to detect in the ear (59).

Differing Levels of Exposure to Notch Ligands Cannot Account for the Choice Between Hair-Cell and Supporting-Cell Fate. Regardless of these uncertainties, our data make one thing clear. The pattern of production of hair cells and supporting cells cannot be determined simply by the pattern of expression of Notch ligands, in the manner proposed by the simple model of lateral inhibition with feedback. The cells that become hair cells are not selected to do so by escape from exposure to Ser1 (they are constantly exposed), Dll1 (its ectopic expression does not change cell fate), or Ser2 (the knockout has only a mild effect). We have shown, however, that hair cells contain Numb, which can block Notch activation (47), supporting the idea that hair cells escape the inhibitory effect of Notch activation not because of lack of ligands from their neighbors, but because they are deaf to the signal delivered by the ligands (see ref. 60, however, for a contrary view of Numb function in the mouse).

We still have to answer our original question: why are hair cells and supporting cells produced in the observed ratio? This cannot be accounted for simply in terms of the rules of asymmetric inheritance of Numb. If each cell in the developing sensory patch went through a final asymmetric division, yielding one daughter that inherited Numb and one daughter that did not, the result would be a 1:1 ratio of hair cells to supporting cells, whereas the measured ratio (in chick basilar papilla) ranges from 1:1.7 to 1:3.9 (18). The level of Numb in the prospective hair cells as opposed to supporting cells may be controlled in some more complex way or through more complex sequences of cell divisions, or some molecule other than Numb and its asymmetrically located companion proteins (61, 63) may confer immunity to lateral inhibition and serve as the key determinant of cell fate.

It is also possible that some factor interacts with the Notch ligands in cis, making those produced by the prospective hair cells more powerful and those produced by the prospective sensory cells less so. If this modulatory factor were itself negatively regulated by Notch activity, we would recover the pattern-generating mechanism of the original simple model, but with the modulatory factor playing the key role instead of Delta. Other signaling pathways also may be involved. Epidermal growth factor signaling, for example, acts in conjunction with the Notch pathway to control cell fate in developing *Drosophila* sense organs (62).

New Insights Come from Ancient Homologies. Our analysis has highlighted many unanswered questions about the role of Notch signaling in the development of the ear. The data leave no doubt, however, that the Notch pathway has a central role in controlling cell fate in this system. Our findings reveal a more complex picture than initially suspected, but reinforce the parallels between the sensory patches of the vertebrate ear and the sensilla of a fly, adding weight to the argument for homology. Thus in both systems, both the Delta and the Serrate subfamilies of Notch ligands are involved; cell fate choice is not simply dictated by the level of Notch ligand production in neighboring cells; and the cells that escape Notch-mediated inhibition contain high levels of Numb, a factor that can block Notch activity by binding to Notch.

Homologies between insects and vertebrates are commonplace at the molecular level. There are innumerable examples of homologous proteins serving the same molecular functions. It is also true that Notch signaling is important in many different tissues, both in vertebrates and invertebrates. But the homologies we have pointed out for the mechanosensory structures go deeper, including function, multicellular anatomy, development, and molecular controls. Indeed, there are few, if any, other multicellular structures where correspondences between insect and vertebrate seem so clear, detailed, and extensive. Of course, there are many differences: the *Drosophila* data can only provide us with hypotheses as to the workings of the vertebrate system, not dogmatic answers. By testing these hypotheses, as we have attempted to do in this paper, we gain a better appreciation of the evolutionary relationship, and a better understanding of the evolutionary relationship opens the way to new hypotheses and new insights into the development of the inner ear.

We are grateful to Yoshio Wakamatsu, Juan-Carlos Izpisua-Belmonte, David Shima, and Guy Richardson for reagents, Jenny Corrigan for cryosectioning, and Helen McNeill for comments on the manuscript. Above all, we are indebted to the colleagues who participated in the earlier work reviewed in this paper: Julie Adam, Catherine Haddon, Domingos Henrique, Christine Hodgetts, David Ish-Horowicz, Yun-Jin Jiang, Alastair Morrison, Anna Myat, and Tanya Whitfield. The Imperial Cancer Research Fund supported us.

1. Watson, G.M., Mire, P. & Hudson, R.R. (1997) *Hearing Res.* **107**, 53–66.
2. Gehring, W.J. & Ikeo, K. (1999) *Trends Genet.* **15**, 371–377.
3. Lewis, J. (1991) *Regeneration of Vertebrate Sensory Cells/Ciba Symp.* **160**, 25–39.
4. Adam, J., Myat, A., Le Roux, I., Eddison, M., Henrique, D., Ish-Horowitz, D. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 4645–4654.
5. Artavanis-Tsakonas, S., Rand, M.D. & Lake, R.J. (1999) *Science* **284**, 770–776.
6. Collier, J.R., Monk, N.A.M., Maini, P.K. & Lewis, J.H. (1996) *J.Theoret. Biol.* **183**, 429–446.
7. Haddon, C., Jiang, Y.-J., Smithers, L. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 4637–4644.
8. Parks, A.L., Huppert, S.S. & Muskavitch, M.A. (1997) *Mech. Dev.* **63**, 61–74.
9. Baker, N.E. (2000) *BioEssays* **22**, 264–273.
10. Guo, M., Jan, L.Y. & Jan, Y.N. (1996) *Neuron* **17**, 27–41.
11. Fekete, D.M. (1996) *Curr. Opin. Neurobiol* **6**, 533–541.
12. Haddon, C. & Lewis, J. (1996) *J.Comp. Neurol* **365**, 113–128.
13. Whitfield, T., Haddon, C. & Lewis, J. (1997) *Semin. Cell Dev. Biol.* **8**, 239–247.
14. Wu, D.K. & Oh, S.H. (1996) *J.Neurosci.* **16**, 6454–6462.
15. Fekete, D.M., Muthukumar, S. & Karagogeos, D. (1998) *J.Neurosci.* **18**, 7811–7821.
16. Stone, J.S. & Rubel, E.W. (1999) *Development (Cambridge, U.K.)* **126**, 961–973.
17. Stone, J.S. & Rubel, E.W. (2000) *J.Comp. Neurol.* **417**, 1–16.
18. Goodyear, R. & Richardson, G. (1997) *J.Neurosci.* **17**, 6289–6301.
19. Corwin, J.T., Jones, J.E., Katayama, A., Kelley, M.W. & Warchol, M.E. (1991) *Ciba Found. Symp.* **160**, 103–130.
20. Kimble, J. & Simpson, P. (1997) *Annu. Rev. Cell Dev. Biol.* **13**, 333–361.
21. Heitzler, P. & Simpson, P. (1991) *Cell* **64**, 1083–1092.
22. Ghysen, A., Dambly-Chaudière, C., Jan, L.Y. & Jan, Y.-N. (1993) *Genes Dev.* **7**, 723–733.
23. Lewis, J. (1996) *Curr. Opin. Neurobiol.* **6**, 3–10.
24. Chitnis, A., Henrique, D., Lewis, J., Ish-Horowitz, D. & Kintner, C. (1995) *Nature (London)* **375**, 761–766.
25. Haddon, C., Smithers, L., Schneider-Maunoury, S., Coche, T., Henrique, D. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 359–370.
26. Henrique, D., Hirsinger, E., Adam, J., Le Roux, I., Pourquie, O., Ish-Horowitz, D. & Lewis, J. (1997) *Curr. Biol.* **7**, 661–670.
27. Morrison, A., Hodgetts, C., Gossler, A., Hrabe de Angelis, M. & Lewis, J. (1999) *Mech. Dev.* **84**, 169–172.
28. Lanford, P.J., Lan, Y., Jiang, R., Lindsell, C., Weinmaster, G., Gridley, T. & Kelley, M.W. (1999) *Nat. Genet.* **21**, 289–292.
29. Riley, B.B., Chiang, M., Farmer, L. & Heck, R. (1999) *Development (Cambridge, U.K.)* **126**, 5669–5678.
30. Haddon, C., Mowbray, C.M., Whitfield, T., Jones, D., Gschmeissner, S. & Lewis, J. (1999) *J.Neurocytol.* **28**, 837–850.
31. Cole, L., Le Roux, I., Nunes, F., Laufer, E., Lewis, J. & Wu, D.K. (2000) *J.Comp. Neurol.* **424**, 509–520.
32. Gho, M., Bellaiche, Y. & Schweisguth, F. (1999) *Development (Cambridge, U.K.)* **126**, 3573–3584.
33. Hartenstein, V. & Posakony, J.W. (1990) *Dev. Biol.* **142**, 13–30.
34. Parks, A.L. & Muskavitch, M.A. (1993) *Dev. Biol.* **157**, 484–496.
35. Walker, R.G., Willingham, A.T. & Zuker, C.S. (2000) *Science* **287**, 2229–2234.
36. Tilney, L.G., Connelly, P., Smith, S. & Guild, G.M. (1996) *J.Cell Biol.* **135**, 1291–1308.
37. Goulding, S.E., White, N.M. & Jarman, A.P. (2000) *Dev. Biol.* **221**, 120–131.
38. Bermingham, N.A., Hassan, B.A., Price, S.D., Vollrath, M.A., Ben-Arie, N., Eatock, R.A., Bellen, H.J., Lysakowski, A. & Zoghbi, H.Y. (1999) *Science* **284**, 1837–1841.
39. Zheng, J.L. & Gao, W.Q. (2000) *Nat. Neurosci.* **3**, 580–586.
40. Bosse, A., Zulch, A., Becker, M.B., Torres, M., Gomez-Skarmeta, J.L., Modolell, J. & Gruss, P. (1997) *Mech. Dev.* **69**, 169–181.
41. Kavalier, J., Fu, W., Duan, H., Noll, M. & Posakony, J.W. (1999) *Development (Cambridge, U.K.)* **126**, 2261–2272.
42. Torres, M., Gomez-Pardo, C. & Gruss, P. (1996) *Development (Cambridge, U.K.)* **122**, 3381–3391.
43. Zeng, C., Younger-Shepherd, S., Jan, L.Y. & Jan, Y.N. (1998) *Genes Dev.* **12**, 1086–1091.
44. Rhyu, M.S., Jan, L.Y. & Jan, Y.N. (1994) *Cell* **76**, 477–491.
45. Wettstein, D.A., Turner, D.L. & Kintner, C. (1997) *Development (Cambridge, U.K.)* **124**, 693–702.
46. Kinsella, T.M. & Nolan, G.P. (1996) *Hum. Gene Ther.* **7**, 1405–1413.
47. Wakamatsu, Y., Maynard, T.M., Jones, S.U. & Weston, J.A. (1999) *Neuron* **23**, 71–81.
48. Bartolami, S., Goodyear, R. & Richardson, G. (1991) *J.Comp. Neurol.* **314**, 777–788.
49. de Celis, J.F. & Bray, S. (1997) *Development (Cambridge, U.K.)* **124**, 241–3251.
50. Lewis, J. (1998) *Semin. Cell Dev. Biol.* **9**, 583–589.
51. Jacobsen, T.L., Brennan, K., Arias, A.M. & Muskavitch, M.A. (1998) *Development (Cambridge, U.K.)* **125**, 4531–4540.
52. Zine, A., Van De Water, T.R. & de Ribaupierre, F. (2000) *Development (Cambridge, U.K.)* **127**, 3373–3383.
53. Irvine, K.D. (1999) *Curr. Opin. Genet. Dev.* **9**, 434–441.
54. Ju, B.G., Jeong, S., Bae, E., Hyun, S., Carroll, S.B., Yim, J. & Kim, J. (2000) *Nature (London)* **405**, 191–195.
55. Zhang, N., Martin, G.V., Kelley, M.W. & Gridley, T. (2000) *Curr. Biol.* **10**, 659–662.
56. Lardelli, M., Dahlstrand, J. & Lendahl, U. (1994) *Mech. Dev.* **46**, 123–136.
57. Hamada, Y., Kadokawa, Y., Okabe, M., Ikawa, M., Coleman, J.R. & Tsujimoto, Y. (1999) *Development (Cambridge, U.K.)* **126**, 3415–3424.
58. Beatus, P., Lundkvist, J., Berg, C. & Lendahl, U. (1999) *Development (Cambridge, U.K.)* **126**, 3925–3935.
59. Myat, A., Henrique, D., Ish-Horowitz, D. & Lewis, J. (1996) *Dev. Biol.* **174**, 233–247.
60. Zhong, W., Jiang, M.M., Schonemann, M.D., Meneses, J.J., Pedersen, R.A., Jan, L.Y. & Jan, Y.N. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 6844–6849.
61. Jan, Y.N. & Jan, L.Y. (2000) *Cell* **100**, 599–602.
62. zur Lage, P. & Jarman, A.P. (1999) *Development (Cambridge, U.K.)* **126**, 3149–3157.
63. Stone, J.S. & Rubel, E.W. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11714–11721.

MOLECULAR GENETICS OF PATTERN FORMATION IN THE INNER EAR: DO COMPARTMENT BOUNDARIES PLAY A ROLE?

John V. Brigande*, Amy E. Kiernan†, Xiaoying Gao*‡, Laurie E. Iten*, and Donna M. Fekete*§

*Department of Biological Sciences, Purdue University, West Lafayette, IN 47907; and †Medical Research Council Institute of Hearing Research, University Park, Nottingham, NG7 2RD, United Kingdom

The membranous labyrinth of the inner ear establishes a precise geometrical topology so that it may subserve the functions of hearing and balance. How this geometry arises from a simple ectodermal placode is under active investigation. The placode invaginates to form the otic cup, which deepens before pinching off to form the otic vesicle. By the vesicle stage many genes expressed in the developing ear have assumed broad, asymmetrical expression domains. We have been exploring the possibility that these domains may reflect developmental compartments that are instrumental in specifying the location and identity of different parts of the ear. The boundaries between compartments are proposed to be the site of inductive interactions required for this specification. Our work has shown that sensory organs and the endolymphatic duct each arise near the boundaries of broader gene expression domains, lending support to this idea. A further prediction of the model, that the compartment boundaries will also represent lineage-restriction compartments, is supported in part by fate mapping the otic cup. Our data suggest that two lineage-restriction boundaries intersect at the dorsal pole of the otocyst, a convergence that may be critical for the specification of endolymphatic duct outgrowth. We speculate that the patterning information necessary to establish these two orthogonal boundaries may emanate, in part, from the hindbrain. The compartment boundary model of ear development now needs to be tested through a variety of experimental perturbations, such as the removal of boundaries, the generation of ectopic boundaries, and/or changes in compartment identity.

Inner Ear Morphogenesis and Gene Expression Patterns

The generation of the vertebrate inner ear requires coordination between morphogenesis and cell fate specification. Morphogenesis will convert a flat epithelial patch (the otic placode) into a hollow sphere (the otic vesicle or otocyst) and finally into a labyrinth of cavities, ducts, and tubules (Fig. 1). It is accompanied by an enormous increase in cell numbers and overall size of the ear, as well focal regions of programmed cell death (for discussion see ref. 1). Cell fate specification will give rise to the neurons of the associated sensory ganglia, the sensory organs for hearing and balance, and nonsensory tissues with various degrees of specialization (2, 3). During the course of these changes, more than 40 different genes are expressed in the developing ear, often in spatially and temporally complex patterns (4, 5). How does this intricate patterning of genes and tissues come about? Can we define some of the rules that govern pattern formation of the inner ear so as to provide a framework for understanding what goes wrong with mutations that lead to congenital defects in the ear?

Several reviews have recently summarized the genes that are expressed early during inner ear development, primarily in mouse, chicken, zebra fish, and *Xenopus* (4–7). With the exception of those that appear to mark the incipient sensory organs, by the otic vesicle stage most of these genes have relatively broad expression domains, although they rarely encompass the entire otic epithelium. That is, they are usually expressed asymmetrically. Highly schematic views of some of these expression patterns are indicated in Fig. 2. Here, we raise the possibility that the domains may abut each other rather abruptly, giving rise to precise boundaries, although we emphasize that this has rarely been rigorously demonstrated. In fact, data from our lab are consistent with some abrupt boundaries, while also showing evidence for partially overlapping expression domains (see below). Gene overlap has also been noted at the ventral pole of the mouse otocyst on the basis of the expression of two transcription factor genes, *Nkx5.1* (*Hmx3*) and *Pax2*, and one receptor tyrosine kinase gene, *EphA4* (formerly called *sek1*) (8). A separate study reported that three members of the Iroquois family of transcription factors are partially overlapping in this same region, at least transiently (9). Clearly, the number and location of the various gene expression domains will continue to undergo further refinement.

Compartment and Boundary Model of Inner Ear Development

Recently, we proposed a model for a global underlying mechanism for pattern formation in the ear that was based on its segregation into separate compartments (2, 10). The model can be stated succinctly as follows:

- (i) The ear is segregated into compartments defined by broad gene expression domains, and it is predicted that these will represent lineage-restriction compartments.
- (ii) Compartment boundaries may specify the location of the sensory organs, endolymphatic duct outgrowth, or emigration of neuronal precursors, by means of short-range signals sent across the boundaries.
- (iii) Compartment identity, perhaps defined by a combination of transcription factors, will specify the structural parts of the ear (semicircular canal, utricle, saccule, cochlea, endolymphatic duct).
- (iv) Compartment identity will also dictate the type of sensory organ (crista, macula, organ of Corti) that forms within each compartment.

Using the *Drosophila* wing and leg imaginal discs as inspiration (11, 12), we reasoned that the compartment boundaries of the ear could be used to pinpoint specific places within this three-dimensional space where key signaling events would occur. In a direct analogy, we would expect the compartment boundaries to also be lineage restriction compartments that resist cell mixing (13). The compartment boundaries would allow for short-range signaling molecules, such as diffusible factors or cell surface

‡Present address: Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285.

§To whom reprint requests should be addressed. E-mail: dfekete@purdue.edu.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: M-L, medial-lateral; A-P, anterior-posterior.

molecules, to signal between the cells of adjacent compartments (14, 15). To take the analogy further, we might expect the result of this inductive signaling to be the secretion of a morphogen or other longer-range signaling molecule that could direct pattern formation on either side of the boundary. However, it should be noted that the distance over which diffusion is likely to act has been difficult to assess in vertebrate systems, and it may be on the order of only a couple hundred micrometers (16). In the ear, the consequences of either short- or long-range inductive signaling might be to dictate the positions of the sensory organs or the outgrowth of the endolymphatic duct, for example. This is shown schematically in Fig. 3.

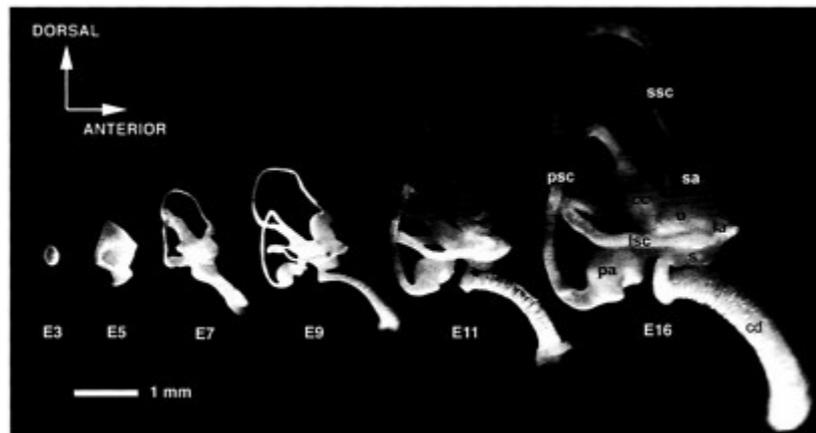


Fig. 1. Morphogenesis of the chicken inner ear viewed by filling the inner ears with opaque paint. Lateral views are shown. The labyrinth develops from a simple ovoid vesicle. Between embryonic day 3 (E3) and E7, the major structural parts of the ear make their appearance, whereas growth, cell fate specification, and tissue differentiation continue for a longer period, cc, Common crus; la, lateral ampulla; lsc, lateral semicircular canal; psc, posterior semicircular canal; sa, superior ampulla; ssc, superior semicircular canal; u, utricle. Modified from Bissonnette and Fekete (58). Both lateral and posterior views of paint-filled ears have been morphed with Elastic Reality (version 1.3; now owned by Avid) to generate a QuickTime movie that can be found as supplementary material on the PNAS website, www.pnas.org.

Note that the model predicts that morphologically elongated sensory organs would require an intersection of only two compartments to define their spatial location. Such organs would include the organ of Corti (in mammals), the basilar papilla (in reptiles and birds), or the vestibular maculae. In contrast, morphologically punctate sensory organs, such as the cristae,

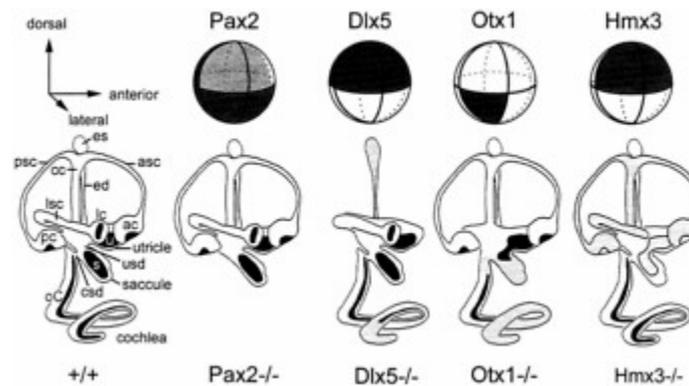


Fig. 2. Schematic representation of the mouse otic vesicle, showing the expression domains of several genes (top row) and the phenotypes that result when the gene is knocked out (bottom row). Dark lines encircling the vesicle indicate theoretical boundaries that segregate the vesicle into compartments. In the bottom row, structures reported missing in the knockout are left off the drawings, while those that have been reported to have a variable phenotype either by a single group or by two different groups are shaded gray. See text for discussion. Data were derived from the following sources: *Pax2* expression, ref. 22; *pax2*^{-/-}, ref. 23; *Hmx3* expression, refs. 8 and 59; *Hmx3*^{-/-}, refs. 20 and 53; *Otx1* expression, refs. 55 and 57; *Otx1*^{-/-}, refs. 54, 55, and 60; *Dlx5* expression and *Dlx5*^{-/-}, ref. 52. asc, Anterior semicircular canal; ac, anterior crista; cc, common crus; csd, cochlear-saccular duct; ed, endolymphatic duct; es, endolymphatic sac; lc, lateral crista; lsc, lateral semicircular canal; oC, organ of Corti; pc, posterior crista; psc, posterior semicircular canal; s, saccular macula; u, utricular macula; usd, utricular-saccular duct.

would require an intersection of at least three compartments to pinpoint location. Note also that the model allows two sensory organs to form adjacent to one other, but separated by a boundary. This review will summarize the current evidence in favor of the compartment boundary model, and suggest future experimental approaches that might offer tests of the model's validity.

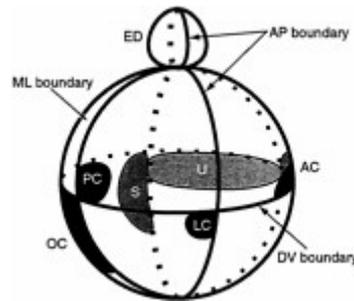


Fig. 3. Model showing one possible arrangement of sensory primordia with respect to the theoretical compartment boundaries defined in Fig. 2. Both chicken and mouse data were used to derive the model, which is viewed from the side with the same orientation as Fig. 2. Several sensory organ primordia arise on the edge of a *SOHo* boundary (10). The lateral crista primordium is located on the edge of the *Otx2* domain (55). The endolymphatic duct also arises on one side of the medial-lateral (ML) boundary at the dorsal pole of the otocyst (see text and Fig. 5). AC, anterior crista; ED, endolymphatic duct; LC, lateral crista; OC, organ of Corti; PC, posterior crista; S, saccular macula; U, utricle.

Sensory Organs Arise at the Boundaries of Broader Gene Expression Domains

One hypothesis suggested by the compartment boundary model is that sensory organs will arise at or adjacent to the boundaries defined by large gene expression domains (Fig. 3). This possibility is now supported by circumstantial evidence in the chicken otocyst. *In situ* hybridization was used to map the spatial domains of two homeobox-containing genes, *SOHo-1* (hereafter called *SOHo*) and *GH6*. These two chicken genes were among the first members of the *Hmx* (*Nkx5*) family of putative transcription factor genes that were reported to be expressed in the developing ear (17, 18). Others include *Hmx2* (*Nkx5.2*) in the mouse and *Hmx3* (*Nkx5.1*) in the mouse and chicken (8, 19, 20). *SOHo* was initially thought to be expressed exclusively in nonsensory tissues of the ear (17), but with the discovery of *Bmp4* as a marker of incipient sensory organs (21), this is now thought to be incorrect. To fully explore the expression domains of *SOHo* and *GH6*, and their relationship to *Bmp4*, we teamed up with F. Nunes and D.K.Wu to hybridize adjacent sections through ears at various developmental stages. The results showed that the two *Hmx* genes seemed to define a lateral "compartment," eventually becoming confined to the territory of the forming semicircular canals (10). At an earlier stage of morphogenesis, several sensory organs formed at the boundaries of this lateral compartment, either just within it (two cristae of the semicircular canals) or just outside of it (the utricle and the basilar papilla/lagena anlagen). Fig. 4 shows the expression patterns of *SOHo* and *Bmp4* at a stage of development when only the anlagen of the anterior and posterior cristae are expressing *Bmp4*. Note that both of these sensory anlagen share a boundary with the *SOHo* expression domain (arrows, Fig. 4). The expression of *SOHo* is somewhat weaker in the *Bmp4*-expressing loci, possibly a prelude to the later stages of sensory organ development, when *SOHo* expression is lost altogether in the hair cell-bearing patches (10).

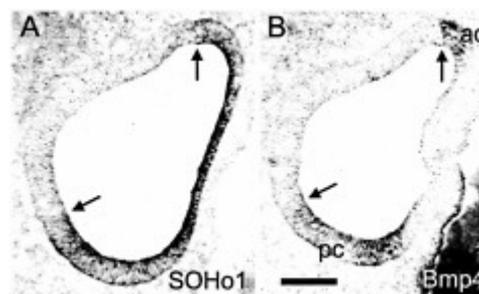


Fig. 4. Sensory organ primordia first arise adjacent to the boundary of a gene, *SOHo-1*, that identifies a putative lateral compartment. Adjacent horizontal sections through the center of the stage 19 chicken otocyst were probed for *SOHo1* mRNA (A) or *Bmp4* mRNA (B). At this stage, *Bmp4* marks the sensory anlagen of the anterior crista (ac) and the posterior crista (pc) (21). Arrows indicate the places where the two genes appear to share the same boundary. Lateral is to the left and anterior is up. (Scale bar=100 μ m.)

The Endolymphatic Duct Arises at a Putative Medial-Lateral (M-L) Boundary

While *SOHo* and *GH6* appear to identify a lateral compartment in the chick, there is evidence that *GATA3*, an unrelated transcription factor, maps to a similar place in the mouse otocyst (M.Holley, personal communication). One notable feature of the three genes, *SOHo*, *GH6*, and *GATA3*, is that they each stop just beyond the base of the endolymphatic duct (Fig. 5 A and C). We were interested in comparing these to medial genes to ascertain whether they also ended abruptly at the dorsal pole, perhaps forming a boundary with the lateral genes. We mapped two molecules in the chick that were interesting in this regard, *PAX2* protein and *EphA4* mRNA. The messages for both genes had previously been described as being expressed in the medial part of the chick or mouse otocyst (8, 22, 23). We have used a commercial antibody against *PAX2* to map the expression of the protein in the ear of the chicken embryo at stages 17–26 (24). Alternate sections were subjected to *in situ* hybridization using probes to *SOHo* or *EphA4*. The results (Fig. 5 B, D, and E) show that *PAX2* and *EphA4* appear to map to the same locations: the medial wall of the vesicle, including the endolymphatic duct, and nearly all of the enlarging cochlear duct region (24). Interestingly, the endolymphatic duct arises at the limit of the *PAX2/EphA4* expression domain at the dorsal pole of the vesicle. Immediately adjacent to this domain, but later separated by a small gap, *SOHo* expression begins. That is, our data indicate that the endolymphatic duct arises just medial to a M-L boundary defined by the juxtaposition of *PAX2/EphA4* and *SOHo* near the dorsal pole (arrows, Fig. 5).

Although it was initially reported that *Pax2* mRNA was not expressed in the endolymphatic duct of the mouse (8, 22), this finding has recently been called into question. From studies using *PAX2* antibodies (M.Holley, personal communication) or *Pax2* mRNA probes (D.K.Wu, personal communication), it now appears that both the protein and the mRNA are detected in the endolymphatic duct at early stages of duct outgrowth in the mouse. Thus, in both the mouse and the chick, *PAX2* defines a medial compartment that includes the endolymphatic duct. Furthermore, this compartment abuts a lateral gene expression domain at the dorsal pole of the otocyst, thus providing circumstantial support in favor of a compartment boundary. It is intriguing that a key morphological structure, the endolymphatic

duct, makes its appearance immediately adjacent to this proposed compartment boundary.

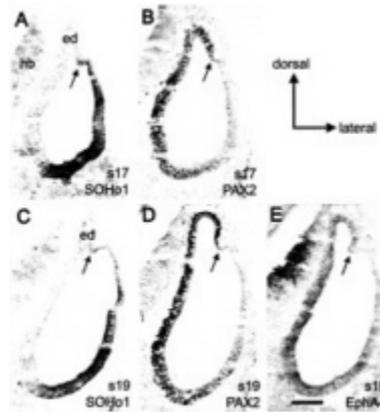


Fig. 5. A putative compartment boundary separating medial and lateral gene expression domains can be found near the dorsal pole of the otocyst in the chicken embryo. Adjacent transverse sections through the middle of the chicken otocyst were probed for *SOHo1* mRNA (A and C), *PAX2* protein (B and D), or *EphA4* mRNA (E). The expression of *EphA4* is generally weak, but appears to map to the same location as *PAX2* at stage 19 (s19), as well as all stages from 18 to 27 (data not shown). Note that all three genes appear to overlap in the ventral but not the dorsal otocyst. Arrows indicate the position of a putative M-L boundary just lateral to the developing endolymphatic duct (ed). (Scale bar=100 μ m.) hb, Hindbrain.

While *PAX2/EphA4* and *SOHo* never seem to overlap in the dorsal part of the otocyst, they clearly overlap in the ventral pole (Fig. 5) (24). This is yet another example of partially overlapping domains in the ventral vesicle, leading to the conclusion that unraveling the number, identity, and timing of gene expression compartments in this part of the vesicle will be a challenge.

Lineage-Restriction Compartments in the Dorsal Otocyst Revealed by Fate Mapping the Chicken Otic Cup

To reveal whether gene expression compartments also represent lineage-restriction compartments, one can mark individual clones of cells and determine whether they mingle freely with unrelated clones. If relatively large clones of cells generate abrupt, linear borders, then a lineage compartment is indicated. Although clonal marking is ideal for addressing this issue, fate mapping slightly larger pools of cells can also reveal lineage compartments. In addition to being useful in detecting boundaries, fate mapping can also provide information about the range of cell movements that accompany morphogenesis. In the ear, this is particularly critical for understanding the topology of otic vesicle closure and how this might aid in the establishment of axial polarity in the otocyst. Without fate mapping, it is difficult to envision how the three major axes of the ear (anterior-posterior, dorsal-ventral, and medial-lateral) arise from a two-dimensional palette, the otic placode. The process of otic cup closure appears to cinch together the outer edges, or rim, of the cup through a purse-string type of closure mechanism. The question arises as to how cells move to accomplish otic cup closure, and how genes with asymmetric expression patterns at the cup stage might be related through these stages of development. That is, it is worthwhile to consider whether a gene expressed at both stages is marking a stable pool of progenitor cells, or whether some cells are down-regulating while others are up-regulating a particular gene. This knowledge is critical for understanding how these stages (placode to vesicle) are related in terms of positional information.

To address the question of what morphogenetic movements of cells accompany cup closure, and to gain further insight into the possible generation of compartments and boundaries, we have performed fate mapping of the cells located on the rim of the otic cup. This was accomplished by targeting small focal injections of fluorescent carbocyanine dyes to the rim of the otic cup. The location of the labeled cells was assessed 24 h later, after cup closure, and 48 h later, in the early stages of morphogenesis. Although the data will be presented in detail elsewhere (44), a schematic summary of the pertinent results is shown in Fig. 6. Fate mapping the otic cup revealed that the dorsal half of the rim gave rise to the endolymphatic duct. Surprisingly, there appeared to be a rather abrupt border that bisected the endolymphatic duct, whereby cells arising from the anterior half did not mix with cells arising from the posterior half. These data suggested that there may be an A-P lineage restriction boundary at this location (Fig. 6). At the present time, there have been no reports of genes whose expression domains clearly respect this putative A-P boundary, bisecting the endolymphatic duct.

Cells from the posteroventral third of the cup gave rise to the entire lateral half of the closed vesicle. Much of this region will go on to form the three semicircular canals. At the time of vesicle closure, these cells came into contact with those that will form the base of the endolymphatic duct, but did not appear to mingle with them. That is, another lineage-restriction boundary is suggested that would prevent the medial (dorsally derived) and lateral (posteroventrally derived) progenitors from mixing. We call this the M-L boundary (Fig. 6). This cell lineage boundary is located at the same place as the gene expression boundary defined by *SOHo* and *PAX2* (arrows, Fig. 5). Thus, the M-L boundary would appear to fulfil both the gene expression and the lineage-restriction criteria needed to represent an authentic compartment boundary.

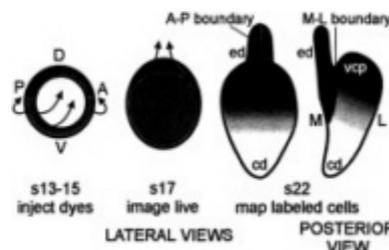


Fig. 6. A schematic of the results of fate mapping the chick otic cup. The data suggest that two lineage-restriction boundaries intersect at the dorsal pole of the otocyst. The M-L lineage-restriction boundary corresponds to the same location as the boundary defined by *SOHo1* and *PAX2* expression domains (compare with Fig. 5). The anterior-posterior (A-P) lineage-restriction boundary appears to bisect the endolymphatic duct (ed). Arrows indicate directions of growth seen at subsequent stages. This schematic is based on data from ref. 44. D, dorsal; V, ventral; cd, cochlear duct; s, stage; vcp, vertical canal pouch.

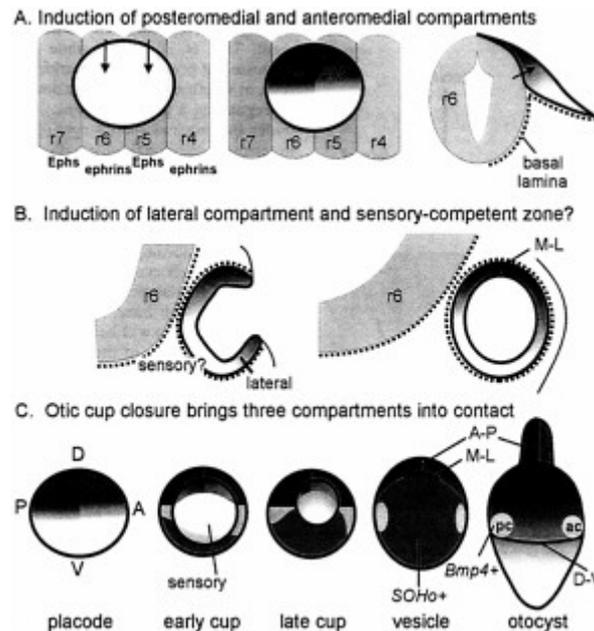


Fig. 7. A schematic showing possible inductive influences of hindbrain on the development of putative compartments in the otic vesicle. (A) The r5-r6 boundary in the hindbrain is aligned with the A-P lineage-restriction boundary in the otic cup. Thus, r5 and r6 are appropriately positioned to send different inductive signals (arrows) to anterior and posterior otic placode, respectively. Because only the dorsal half of the placode makes intimate contact with the hindbrain cells because of the absence of a basal lamina between the two structures (34), inductive signals that require direct cell contact may not extend to the ventral rim of the otic field, as shown. (B) It is proposed that the ventral tissue does not acquire lateral identity until the cup stage, and is separated from the dorsomedial tissue by a region that acquires sensory competence. The signaling sources for either lateral or sensory identity are unknown, but they could be influenced by distance from the hindbrain and/or proximity to lateral ectoderm or mesoderm. (C) A schematic of otic vesicle formation, combining information from fate mapping and gene expression domains of *SOHo* and *Bmp4*. The gray region in the center of the field is proposed to correspond to a sensory-competent region that will intersect with the broader gene expression domains to form the sensory patches and perhaps the ganglion cells. Only the formation of the anterior crista (ac) and posterior crista (pc) is shown. The timing of appearance and location of a putative dorsal-ventral (D-V) boundary, although predicted to be present by stage 22 to place the cristae at the right location, has not yet been confirmed by fate mapping, cd, Cochlear duct; ed, endolymphatic duct; r, rhombomere.

Three Compartments Appear to Intersect at the Dorsal Pole of the Vesicle

As predicted by the compartment-boundary model, there appears to be an intersection of three compartments (anteromedial, posteromedial, and lateral) at the dorsal pole of the vesicle that is ideally positioned to provide the information needed to specify the location of endolymphatic duct outgrowth. We predict that the lateral compartment may be sending an inductive signal to the two adjacent compartments, which then act or interact together to generate another signal directing cells near the boundary intersection to evaginate dorsally. The lateral compartment is not drawn up into this elongating duct, but remains a distinct compartment.

Within a few stages, the lateral compartment will independently begin to enlarge dorsally to form the vertical canal pouch (Fig. 6, vcp), which eventually becomes the anterior and posterior canals. Perhaps canal outgrowth is controlled by a signal sent in the opposite direction to those specifying duct outgrowth. That is, a signal might originate in both medial compartments that crosses all along the M-L boundary to direct canal outgrowth as a thin plate oriented in the A-P direction. The formation of the vertical pouch causes the former "dorsal" pole of the vesicle (where A-P and M-L boundaries originally intersected) to lie in a deepening crevasse between the endolymphatic duct and the vertical canal pouch.

Is the Hindbrain a Source of Patterning Information for Inner Ear A-P Compartments?

If, as suggested by the fate mapping, there are two intersecting compartment boundaries formed as the cup closes, then the implication is that the cells constituting the three compartments must have distinct identities that keep them from mixing. One possibility is that the ear first acquires general axial polarity, such as A-P or dorsal-ventral polarity. Because the ear is a relatively late-developing embryonic field, it should be possible for it to obtain information about its axial polarity from adjacent tissues. This positional information could take the form of broad gradients, or it could be imposed more directly as absolute differences between neighboring cells, depending on the source(s) of the putative inductive signal(s). An ideal candidate for a source of positional information is the hindbrain (25–30). Consider first how the hindbrain might contribute information to specify the anterior and posterior lineage-restriction compart

ments that appear to bisect the dorsomedial otocyst. It is noteworthy that the otic cup forms adjacent to two hindbrain compartments, rhombomeres 5 and 6 (r5 and r6). The boundary between the two rhombomeres is aligned almost precisely with the middle of the otic cup. Therefore, the A-P boundary of the otic cup is aligned with the r5-r6 boundary (Fig. 7A). This seems unlikely to be a chance arrangement. The r5-r6 boundary is one of the earliest to make its appearance; it is formed by early stage 9 (31), when the otic placode is just beginning to form (32–34). We speculate that r5 and r6 may be sending different signals to the anterior and posterior otocyst, respectively. We further suggest that these signals might be transferred most effectively or even exclusively to the dorsal half of the vesicle, where the neural tube and otic placode cells come into direct contact because of the absence of a basal lamina between them (34). This would generate, from the start, separate anteromedial (AM) and posteromedial (PM) compartments that are detectable by fate mapping (Fig. 7A).

One cell-cell signaling system that could potentially be involved in sending differential positional information from r5 and r6 to the adjacent otic field is the Eph/ephrin system (35, 36). r6 cells express the ligands, ephrin B2 and B3, at high levels but express the receptor tyrosine kinase, EphA7, only weakly; in contrast, r5 cells do not express the ephrins but express the receptors, EphA4, -A7, -B2, and -B3 at high levels (28, 37, 38). Thus, the posteromedial otic cells in direct contact with r6 are positioned to receive an ephrin-mediated signal that those adjacent to r5 would not. Do the otic placode cells express the appropriate receptors? To date, there are no systematic studies at the appropriate stages of development. By the vesicle stage, the medial side expresses at least two receptors, EphA4 and EphA7 (8, 28). One of these, EphA4, is reportedly able to bind not only to the ephrin A ligands but also to ephrin B2 and B3 (37, 39), which are expressed by r6. Clearly, more information is needed about the timing and full range of Eph/ephrin expression in the otic cup to fully explore this potential signaling system. Of course, other types of signals, including diffusible molecules such as bone morphogenic proteins (BMPs) are also candidates for inductive interactions between neural tube and otic epithelium.

The proposed requirement of r5 and/or r6 for inducing A-P differences in the otocyst raises an interesting prediction: that the A-P boundary of the dorsomedial otocyst would be compromised in mutants where r5 and/or r6 fails to form or where the otic epithelium arises far from the hindbrain. According to our model, defective outgrowth of the endolymphatic duct should be one result of all such mutants. Mutants that fulfill one or both of these criteria and are missing the endolymphatic duct include knockouts of *Hoxal* (26, 40–43) and the *kreisler* mouse mutant (27–29, 45). Mice lacking two retinoic acid receptors, *RAR α* and *RAR β* , provide an informative counterexample (46). These mice have defects in hindbrain patterning near the otic field, including an expanded r5 and a combined r6/7 rhombomere. However, the r5/r6 boundary is still present with an expanded otic placode arising next to it. Despite the alterations in r5 and r6, the otic vesicle closes, the endolymphatic duct forms, and the inner ear displays no abnormalities by 14.5 days post coitus. This may be explained by the presence of the r5/r6 boundary at the otic placode stage.

Consider next the medial versus lateral positional information. Again it seems plausible for the neural tube cells to interact directly with the medial placode, but not the lateral placode, to provide it with general medial positional information. One possibility is that the lateral cells acquire a distinct identity from the medial cells, either because they do not receive a medializing signal from the neural tube or because they fail to receive an inhibitory signal from the neural tube that prevents them from becoming lateral. However, we suspect that specification of M-L identity may not be this simple, and instead speculate that a lateral compartment may arise later and spatially nonadjacent to the dorsomedial compartments at the placode stage (Fig. 7B). There are several reasons for making this suggestion. First, if a lateral compartment (L) were to form immediately adjacent to the two dorsomedial compartments (AM and PM), then the three compartments would already intersect as early as the placode stage, in the center of the placode. This situation is both temporally and spatially inappropriate for the specification of endolymphatic duct outgrowth according to our model. Second, if we use *SOHo* as an indicator of lateral identity, we find that it is not expressed until the otic cup stage in the chick, and then only along the posteroventral rim of the cup that is destined to become the lateral compartment (10). Third, there is additional evidence from gene expression data that the center (the ventromedial part) of the invaginating cup may differ from surrounding tissues in the expression of *Lmx1*, a gene whose expression levels are down-regulated by the adjacent ventral (but not dorsal) neural tube (47). So we are suggesting that lateral identity is acquired or becomes fixed slightly later than medial identity, and only by the lateral-most part of the invaginating cup (as shown in Fig. 7B). The nature of such signal(s) is unclear; it may involve specification and/or maintenance from the adjacent ectoderm. Assuming that M-L differences do arise by the otic cup stage, we would predict that these compartments would differ in their adhesive properties. In this way, when the two groups of cells first meet at the dorsal pole of the otocyst at the time of cup closure, they would resist mixing.

Perturbations that cause abnormal separation between the neural tube and the otic placode would be predicted to cause insufficient or abnormal M-L distinctions within the otic epithelium. In this regard, it is interesting to consider the *kreisler* mutant phenotype (27–30, 45). In these mice, the otic cup develops abnormally far from the hindbrain, and the expression domain of at least one otic gene, *Fgf3*, is likewise displaced from lateral to medial (48). As a corollary, endolymphatic duct outgrowth is often severely truncated or absent. A second example is the case of the retinaldehyde dehydrogenase 2 knockout mice (*Raldh2^{-/-}*) which are unable to synthesize retinoic acid. This defect causes severe alterations in hindbrain patterning and gene expression (49, 50), including a failure of r5 and r6 to be properly specified. The hindbrain defects are accompanied by abnormal expression of M-L genes in the otocyst: *Pax2* transcripts are absent and *Nkx5.1* (*Hmx3*) transcripts are expressed throughout the otocyst at 9.5 days post coitus (dpc). This suggests that the otocyst is completely lateralized and the M-L boundary is missing. The endolymphatic duct does not appear to form by 10.5 dpc, when the *Raldh2^{-/-}* mice die. The *Raldh2^{-/-}* phenotype suggests that the hindbrain participates in establishing asymmetric M-L gene expression in the otocyst, although a direct effect of retinoids on the ear (51) cannot be ruled out.

Testing a Compartment Boundary Model of Ear Morphogenesis

Ultimately, the utility of a model lies in its predictive value. In the case of the inner ear, it should be possible to predict how the system will respond to different kinds of perturbations, including those that should disrupt compartment boundaries, or those that would be expected to change compartment identity. The generation of knockout mice lacking one or more of the genes expressed in the inner ear is one experimental approach (6). If the genes showing broad expression domains are providing information to specify the identity of the compartments, then eliminating an individual gene might be equivalent to changing the compartment identity. In addition to this, or as an alternative consequence, the gene knockout might disrupt or eliminate one or more of the compartment boundaries defined by the gene of interest. One can look carefully at the phenotypes (shown schematically in Fig. 2B) and try to make sense of them in light of a compartment-boundary model. The phenotypes generated by *Pax2* and *Dlx5* seem to come closest to suggesting that these two genes each specify a compartment that will develop into the cochlea or the vertical canals, respectively. However, this interpretation is complicated by the fact that the expression domain of *Pax2* (and possibly *Dlx5*) includes tissue fated to form the endolymphatic duct in the chick (J.V.B and D.M.F., unpublished observations), and yet the duct is not always reported to be defective in the mutants (23, 52).

The lateral canal and its sensory organ are of particular interest with respect to the mutant phenotypes. The crista ampullaris has been reported missing in both *Hmx3* and *Otx1* knockout mice (53, 54). In the schematic, these two genes are mapped to adjacent compartments, and it is known that the lateral crista arises within the *Otx1* domain at its dorsal border (55). Although the nonoverlapping domains of *Hmx3* and *Otx1* must be verified by careful study of adjacent sections, it remains possible that interfering with the expression domains on either side of a compartment boundary can result in the specific loss of the lateral crista. This result would be predicted by the compartment boundary model, assuming that the boundary could not be maintained in the absence of either one of the two genes.

Our hypothesis is that the convergence of three compartments, AM, PM, and L, is a necessary prerequisite for endolymphatic duct outgrowth; one of the most obvious tests would be to generate an ectopic convergence of these boundaries elsewhere in the otic epithelium. The model makes a very specific prediction about what should happen if an ectopic L compartment were created on the A-P boundary in the medial domain: two extra (supernumerary) endolymphatic ducts should arise in this instance. It will be a challenge to perform such an experiment, in view of the fact that some genes appear to regulate (e.g., *Hmx3*) whereas others (e.g., *Pax2*) do not (20, 56, 57) in situations where the otic field is rotated or transplanted ectopically. Nonetheless, if properly executed, the result of creating ectopic compartment boundaries could well prove to be as intriguing, and enlightening, as the classical studies in developing and regenerating limbs.

We thank J.P. Bissonnette for the paint-fill images used to prepare Fig. 1 and the ear morph movie. Contact D.M.F. to download a higher-resolution version of the movie. We thank Connie Cepko for providing probes. This work was supported by grants from the National Institutes of Health (to D.M.F.), the March of Dimes Birth Defects Foundation (to D.M.F.), and the National Organization for Hearing Research (to J.V.B.).

1. Lang, H., Bever, M.M. & Fekete, D.M. (2000) *J. Comp. Neurol.* **417**, 205–220.
2. Fekete, D.M. (1996) *Curr. Opin. Neurobiol.* **6**, 533–541.
3. Whitfield, T., Haddon, C. & Lewis, J. (1997) *Semin. Cell Dev. Biol.* **8**, 239–247.
4. Torres, M. & Giraldez, F. (1998) *Mech. Dev.* **71**, 5–21.
5. Rivolta, M.N. (1997) *Audiol. Neurootol.* **2**, 36–49.
6. Fekete, D.M. (1999) *Trends Neurosci.* **22**, 263–269.
7. Ryan, A.F. (1997) *Semin. Cell Dev. Biol.* **8**, 249–256.
8. Rinkwitz-Brandt, S., Arnold, H.H. & Bober, E. (1996) *Hear. Res.* **99**, 129–138.
9. Bosse, A., Zülch, A., Becker, M.-B., Torres, M., Gomez-Skarmeta, J.L., Modellell, J. & Gruss, P. (1997) *Mech. Dev.* **69**, 169–181.
10. Kiernan, A. E., Nunes, F., Wu, D.K. & Fekete, D.M. (1997) *Dev. Biol.* **191**, 215–229.
11. Meinhardt, H. (1983) *J. Embryol. Exp. Morphol.* **76**, 115–137.
12. Meinhardt, H. (1983) *Dev. Biol.* **96**, 375–385.
13. Dahmann, C. & Basler, K. (1999) *Trends Genet.* **15**, 320–326.
14. Lawrence, P.A. & Struhl, G. (1996) *Cell* **85**, 951–961.
15. Shubin, N., Tabin, C. & Carroll, S. (1997) *Nature (London)* **388**, 639–648.
16. Yang, Y., Drossopoulou, G., Chuang, P.T., Duprez, D., Marti, E., Bumcrot, D., Vargesson, N., Clarke, J., Niswander, L., McMahon, A., *et al.* (1997) *Development (Cambridge, U.K.)* **124**, 4393–4404.
17. Deitcher, D. L., Fekete, D.M. & Cepko, C.L. (1994) *J. Neurosci.* **14**, 486–498.
18. Stadler, H.S. & Solursh, M. (1994) *Dev. Biol.* **161**, 251–262.
19. Bober, E., Baum, C., Braun, T. & Arnold, H.H. (1994) *Dev. Biol.* **162**, 288–303.
20. Herbrand, H., Guthrie, S., Hadrys, T., Hoffmann, S., Arnold, H., Rinkwitz-Brandt, S. & Bober, E. (1998) *Development (Cambridge, U.K.)* **125**, 645–654.
21. Wu, D.K. & Oh, S.-H. (1996) *J. Neurosci.* **16**, 6454–6462.
22. Nornes, H.O., Dressler, G.R., Knapik, E.W., Deutsch, U. & Gruss, P. (1990) *Development (Cambridge, U.K.)* **109**, 797–809.
23. Torres, M., Gomez-Pardo, E. & Gruss, P. (1996) *Development (Cambridge, U.K.)* **122**, 3381–3391.
24. Fekete, D.M. & Gao, X. (2000) in *Molecular and Cellular Biology of the Ear*, ed. Lim, D. (Plenum, New York), pp. 99–112.
25. Mansour, S.L., Goddard, J.M. & Capecchi, M.R. (1993) *Development (Cambridge, U.K.)* **117**, 13–28.
26. Mark, M., Lufkin, T., Vonesch, J.L., Ruberte, E., Olivo, J.C., Dolle, P., Gorry, P., Lumsden, A. & Chambon, P. (1993) *Development (Cambridge, U.K.)* **119**, 319–338.
27. Deol, M.S. (1964) *J. Embryol. Exp. Morphol.* **12**, 475–490.
28. Manzanares, M., Trainor, P.A., Nonchev, S., Ariza-McNaughton, L., Brodie, J., Gould, A., Marshall, H., Morrison, A., Kwan, C.T., Sham, M.H., *et al.* (1999) *Dev. Biol.* **211**, 220–237.
29. Cordes, S.P. & Barsh, G.S. (1994) *Cell* **79**, 1025–1034.
30. Frohman, M.A., Martin, G.R., Cordes, S.P., Halamek, L.P. & Barsh, G. (1993) *Development (Cambridge, U.K.)* **117**, 925–936.
31. Fraser, S., Keynes, R. & Lumsden, A. (1990) *Nature (London)* **344**, 431–435.
32. Meier, S. (1978) *Anat. Rec.* **191**, 447–458.
33. Hilfer, S.R., Esteves, R.A. & Sanzo, J.F. (1989) *J. Exp. Zool.* **251**, 253–264.
34. Mayordomo, R., Rodriguez-Gallardo, L. & Alvarez, I.S. (1998) *J. Anat.* **193**, 35–48.
35. Xu, Q., Mellitzer, G., Robinson, V. & Wilkinson, D.G. (1999) *Nature (London)* **399**, 267–271.
36. Mellitzer, G., Xu, Q. & Wilkinson, D.G. (2000) *Curr. Opin. Neurobiol.* **10**, 400–408.
37. Bergemann, A.D., Zhang, L., Chiang, M.K., Brambilla, R., Klein, R. & Flanagan, J.G. (1998) *Oncogene* **16**, 471–480.
38. Gale, N.W., Flenniken, A., Compton, D. C., Jenkins, N., Copeland, N.G., Gilbert, D.J., Davis, S., Wilkinson, D.G. & Yancopoulos, G.D. (1996) *Oncogene* **13**, 1343–1352.
39. Gale, N.W., Holland, S.J., Valenzuela, D. M., Flenniken, A., Pan, L., Ryan, T. E., Henkemeyer, M., Strebhardt, K., Hirai, H., Wilkinson, D.G., *et al.* (1996) *Neuron* **17**, 9–19.
40. Carpenter, E.M., Goddard, J.M., Osamu, C., Manley, N.R. & Capecchi, M.R. (1993) *Development (Cambridge, U.K.)* **118**, 1063–1075.
41. Chisaka, O., Musci, T.S. & Capecchi, M.R. (1992) *Nature (London)* **355**, 516–520.
42. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. (1991) *Cell* **66**, 1105–1119.
43. Rossel, M. & Capecchi, M.R. (1999) *Development (Cambridge, U.K.)* **126**, 5027–5040.
44. Brigande, J.V., Iten, L.E. & Fekete, D.M. (2000) *Dev. Biol.*, in press.
45. McKay, I.J., Muchamore, L., Krumlauf, R., Maden, M., Lumsden, A. & Lewis, J. (1994) *Development (Cambridge, U.K.)* **120**, 2199–2211.
46. Dupe, V., Ghyselinck, N.B., Wendling, O., Chambon, P. & Mark, M. (1999) *Development (Cambridge, U.K.)* **126**, 5051–5059.
47. Giraldez, F. (1998) *Dev. Biol.* **203**, 189–200.
48. McKay, I.J., Lewis, J. & Lumsden, A. (1996) *Dev. Biol.* **174**, 370–378.
49. Niederreither, K., Vermot, J., Schuhbauer, B., Chambon, P. & Dolle, P. (2000) *Development (Cambridge, U.K.)* **127**, 75–85.
50. Niederreither, K., Subbarayan, V., Dolle, P. & Chambon, P. (1999) *Nat. Genet.* **21**, 444–448.
51. Choo, D., Sanne, J.L. & Wu, D.K. (1998) *Dev. Biol.* **204**, 136–150.
52. Acampora, D., Merlo, G.R., Paleari, L., Zerega, B., Postiglione, M.P., Mantero, S., Bober, E., Barbieri, O., Simeone, A. & Levi, G. (1999) *Development (Cambridge, U.K.)* **126**, 3795–3809.
53. Wang, W., Van De Water, T. & Lufkin, T. (1998) *Development (Cambridge, U.K.)* **125**, 621–634.
54. Acampora, D., Mazan, S., Avantageggiato, V., Barone, P., Tuorto, F., Lallemand, Y., Brulet, P. & Simeone, A. (1996) *Nat. Genet.* **14**, 218–222.
55. Morsli, H., Tuorto, F., Choo, D., Postiglione, M.P., Simeone, A. & Wu, D.K. (1999) *Development (Cambridge, U.K.)* **126**, 2335–2343.
56. Hutson, M.R., Lewis, J.E., Nguyen-Luu, D., Lindberg, K.H. & Barald, K.F. (1999) *J. Neurocytol.* **28**, 795–807.
57. Wu, D. K., Nunes, F.D. & Choo, D. (1998) *Development (Cambridge, U.K.)* **125**, 11–20.
58. Bissonnette, J.P. & Fekete, D.M. (1996) *J. Comp. Neurol.* **368**, 620–630.
59. Rinkwitz-Brandt, S., Justus, M., Oldenettel, L., Arnold, H.-H. & Bober, E. (1995) *Mech. Dev.* **52**, 371–381.
60. Acampora, D., Avantageggiato, V., Tuorto, F., Barone, P., Reichert, H., Finkelstein, R. & Simeone, A. (1998) *Development (Cambridge, U.K.)* **125**, 1691–1702.

PATTERNING OF THE MAMMALIAN COCHLEA

Raquel Cantos*, Laura K.Cole*, Dario Acampora[†], Antonio Simeone^{†‡}, and Doris K.Wu*[§]

*National Institute on Deafness and Other Communication Disorders, Rockville, MD 20850; [†]International Institute of Genetics and Biophysics, Consiglio Nazionale delle Ricerche, 1–80125 Naples, Italy; and [‡]Medical Research Council Centre for Developmental Neurobiology, Fourth Floor, New Hunt's House, Guy's Campus, King's College London, London Bridge, London SE1 1UL, United Kingdom

The mammalian cochlea is sophisticated in its function and highly organized in its structure. Although the anatomy of this sense organ has been well documented, the molecular mechanisms underlying its development have remained elusive. Information generated from mutant and knockout mice in recent years has increased our understanding of cochlear development and physiology. This article discusses factors important for the development of the inner ear and summarizes cochlear phenotypes of mutant and knockout mice, particularly *Otx* and *Otx2*. We also present data on gross development of the mouse cochlea.

The mammalian cochlea, the end organ of auditory function, is a truly remarkable structure. Its uniquely coiled shape, diversity of cell types, and intricate architecture are unmatched by any other organs in the body. The sensory component of the cochlea, the organ of Corti, consists of both sensory hair cells and supporting cells, and it spirals like a ribbon down the cochlear duct. The cochlea is tonotopically mapped so that the hair cells at each location along the cochlea are most sensitive to a particular frequency (for review, see ref. 1). Multiple structural features of the hair cells are organized in a gradient along the cochlea that could contribute to the differential frequency selectivity of the hair cells (for review, see ref. 2). For example, hair cells in the base of the cochlea have shorter cell bodies and their stereocilia are shorter and more abundant than those of hair cells in the apex. In addition, the width of the basilar membrane and the mass of the tectorial membrane also increase toward the apex of the cochlea. These overall structural gradients along the cochlea are largely conserved among different species but vary depending on the range of absolute frequencies detected and the most sensitive frequency range of an individual species. Little is known about the molecular mechanisms that establish the fine structural patterning of the cochlea or that underlie the tonotopic organization of the organ. Likewise, little is known about what makes a cochlea coil and what dictates the variation in the number of coils among different species (3). Recent gene targeting approaches in mice have provided insights by identifying a number of genes important for the shaping of the cochlea at both the gross and fine structural levels. Here, we summarize data from mutant and knockout mice with cochlear defects and highlight several features of the gross development of the cochlea that may pertain to its mature functions.

Gross Development of the Cochlea

The mouse inner ear can be roughly divided into a dorsal vestibular and a ventral saccular and cochlear region. The cochlea develops from the ventral portion of the rudimentary otocyst. Fig. 1 illustrates a series of developing inner ears in which the lumen has been filled with a latex paint solution to reveal its gross anatomy. At 10.75 dpc (days postcoitum), the cochlear anlage becomes evident, and it first extends ventromedially and then anteriorly. As a result, the first turn of the cochlea is apparent at 12 dpc. Over time, there is a continual increase in length in the proximal region (Fig. 1, arrows) as well as coiling in the distal region of the cochlea to achieve its mature 1.75 turns by 17 dpc (4). Based on birth-dating studies, it has been proposed that the junction of the presumptive saccule and cochlea is the site of cochlear growth (5). Our paint-fill data confirmed that there is indeed a considerable increase in distance between the presumptive saccule and the location of the first turn of the cochlear duct over time (Fig. 1, distance between arrows). However, we cannot verify from these results whether the cochlea grows exclusively at the junction of the saccule and cochlea. Determining how the cochlea grows remains an important question because it may establish the basis for its tonotopic organization.

The development of the vestibular component in the dorsal portion of the otic vesicle occurs concurrently with ventral cochlear development. As morphogenesis progresses, there is a gradual restriction in the portion of the inner ear separating the dorsal and ventral regions. By 15 dpc, the utricle and saccule are separated by the utriculosaccular duct (Fig. 1, *usd*), which is continuous with the endolymphatic duct. By 16.5 dpc, when morphogenesis is more advanced, the utriculosaccular duct becomes progressively restricted, and the utricle and saccule are essentially two separate chambers (Fig. 2). As a result, a single injection of latex paint solution into the saccule, for example, invariably fills only four components of the labyrinth, including the endolymphatic sac and its duct, the saccule, and the cochlea (Fig. 2A). In contrast, single injections to either the utricle or ampulla fill the rest of the vestibular system (Fig. 2B). It is conceivable that the utricle and saccule are connected in the mature mouse inner ear by the utriculosaccular duct, but the lumen is too narrow for the passage of the paint solution. Nevertheless, these results suggest that there are largely two separate chambers in the membranous labyrinth of a mature mouse inner ear, and only the cochlea and the saccule are efficiently under fluid regulation by the endolymphatic sac. This gross anatomical finding is supported by the fact that the endolymph in the ampullae and cochlea have different ionic compositions and electrochemical potentials (6). In addition, a recent report of *EphB2*-knockout mice shows that loss of this member of Eph receptor tyrosine kinase family resulted in vestibular defects that may be associated with reduced endolymph production in the ampulla and canal regions (7). Despite the reduced size of vestibular membranous labyrinth, these mice have no apparent cochlear defects, and the membranous labyrinth of the cochlear region appeared normal (ref. 7 and B.Fritzsch, personal communication). Taken together, these results support the paint-fill data that the membranous labyrinth of a mature inner ear consists of two separate compartments. This difference in the properties of the endolymph within different inner ear components such as the cochlea and ampulla may play a direct role in facilitating the specific functions of each component. Furthermore, such functions may be

[§]To whom reprint requests should be addressed at: National Institute on Deafness and other Communication Disorders, 5 Research Court, Room 2B34, Rockville, MD 20850. E-mail: wud@nidcd.nih.gov.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviation: dpc, days postcoitum.

affected in mutants in which the utricle and saccule fail to form separate chambers (8–11).

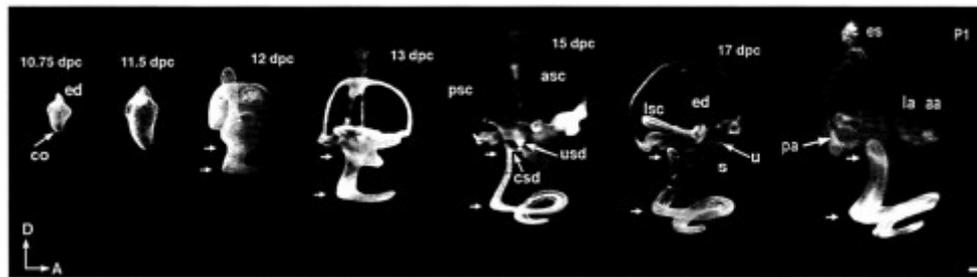


Fig. 1. Lateral views of paint-filled membranous labyrinths. Membranous labyrinths of inner ears from 10.75 dpc to postnatal day 1 were filled with latex paint solution as described. At 10.75 dpc, the protrusions of the endolymphatic duct in the dorsal and the cochlear anlage in the ventral portion of the otocyst are evident. By 17 dpc, the gross anatomy of the inner ear is mature. Arrowheads identify the proximal region of the cochlea, aa, anterior ampulla; asc, anterior semicircular canal; co, cochlea; csd, cochleosaccular duct; ed, endolymphatic duct; es, endolymphatic sac; la, lateral ampulla; lsc, lateral semicircular canal; pa, posterior ampulla; psc, posterior semicircular canal; s, saccule; u, utricle; usd, utriculosaccular duct. Orientation: D, dorsal; A, anterior. (Scale bar=100 μm .) [Adapted with permission from Morsli *et al.* (4) (Copyright 1998, the Society for Neuroscience).]

Unlike the mouse, the endolymphatic duct in humans has a bifurcation connecting to both the utricle and saccule. Therefore, structures such as the ampullae and utricle in humans have a more direct access to fluid regulation by the endolymphatic sac than do these same structures in the mouse. This anatomical difference between humans and mice might be an important consideration when evaluating mouse models for human genetic disorders affecting fluid homeostasis.

Gross Patterning of the Cochlea

The normal development of the inner ear is thought to depend on multiple surrounding tissues, including the hindbrain, neural crest, mesenchyme, and possibly the notochord (for review, see refs. 12–16). Some of the genes involved have been identified, and the examination of mutants and knockout mice demonstrates that the absence of these gene products invariably affects the patterning of the cochlea as well (Table 1).

Genes Expressed in the Otic Epithelium. Several genes expressed in the otic epithelium are important for the normal development of the cochlea (Table 1). For example, the absence of *Pax2*, a paired-box transcription factor, leads to agenesis of the cochlea (17). The development of inner ears of *Eya1* (*eyes absent*)-knockout mice arrest at the otic vesicle stage (18). In these mice, the endolymphatic duct is either absent or malformed, and the VIIIth ganglion fails to form. Both *Otx1* and *Otx2* are expressed in the otocyst and are important for cochlear and vestibular development. As such genetic information accumulates, it is important to determine when and where these genes act along the developmental pathway. This task is complicated by the fact that often several members of a single gene family, which may share redundant functions, are expressed in the inner ear during development. For example, *Pax2* and -8 are both expressed in the otic epithelium (18–20). However, no inner ear phenotypes have been reported for *Pax8*-knockout mice so far (21). *Eya1* and -2 are both expressed in the VIIIth ganglion (22). *Dlx2*, -3 , -5 , and -6 are all expressed in the mouse inner ear (23–26). However, only the knockout of *Dlx5* has been reported to display inner ear defects, including abnormalities in the semicircular canals, ampullae, endolymphatic duct, and cochlea (25–28). Therefore, it is important to sort out specific functions for each member of a gene family. We have attempted to address this issue for *Otx1* and -2 .

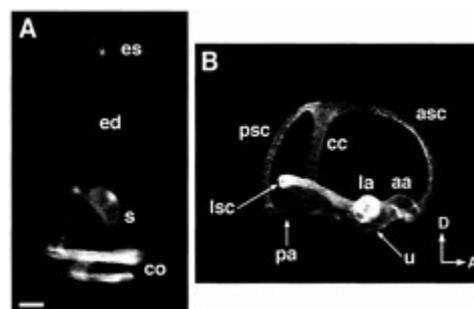


Fig. 2. Partially paint-filled mouse membranous labyrinths at 16.5 dpc. At 16.5 dpc, the membranous labyrinth largely consists of two compartments. (A) Latex paint solution injected into the endolymphatic sac fills the endolymphatic duct, saccule, and cochlea only. (B) Latex paint solution injected into the lateral ampulla fills the utricle, the three semicircular canals, and their ampullae. cc, common crus. Orientation as per Fig. 1. (Scale bar=200 μm .)

Otx Genes. *Otx1* and *Otx2* are murine orthologues of the *Drosophila orthodenticle* gene. These genes are bicoid-like transcription factors important for the development of the head and sense organs (29–32). In the inner ear, both *Otx1* and *Otx2* are activated during the otocyst stage. At 10.25 dpc, *Otx1* is expressed in the ventrolateral wall of the otocyst, and *Otx2* is expressed in the ventral tip of the otocyst within a portion of the *Otx1*-positive region (10). As development progresses, the most dorsal boundary of the *Otx1* domain corresponds to the presumptive lateral canal level, and the dorsal boundary of *Otx2* expression domain corresponds to the middle of the utricular anlage (Fig. 3A, schematic). In *Otx1*-knockout mice, the lateral canal and ampulla are missing, the cochlea is misshapen, and the utricle and saccule fail to form separate chambers (Fig. 3B; refs. 10 and 29). *Otx2*-knockout mice die around 10 dpc before any significant inner ear development (30, 33). To address the possible functions of *Otx2* in inner ear development, inner ears of *Otx1*-knockout mice with one disrupted allele of *Otx2* (*Otx1* $^{-/-}$; *Otx2* $^{+/-}$) were examined. These mice show much more severe defects, particularly in ventral structures, including the saccule and cochlea, which normally express *Otx2* (Fig. 3C and Table 2).

The predicted murine OTX1 and OTX2 proteins show extensive homology between the N terminus and the end of the homeodomain. *otd* and its orthologue, *Otx2*, have been shown to largely

substitute for *Otx1* functions in the brain (34, 35). Therefore, we attempted to determine whether *Otx2* and *otd* could also substitute for *Otx1* in inner ear development, especially in the lateral canal region where normally only *Otx1* is expressed. To achieve this substitution, the full-length human *Otx2* cDNA was introduced into a disrupted *Otx1* locus (*hOtx2¹/hOtx2¹*) and thus placed under transcriptional regulation of *Otx1*. The inner ears of these mice had no lateral canal or ampulla, indicating that *Otx2* was not able to functionally compensate for *Otx1* in forming these structures (refs. 10 and 34; Fig. 3D). In regions where the two genes are normally coexpressed, there was a partial rescue of saccular and cochlear phenotypes, as well as the separation of utricle and saccule (Table 2). Introduction of *otd* into the disrupted *Otx1* locus also failed to restore the formation of the lateral canal and ampulla (ref. 35; Fig. 3E). In addition, although the shape of the cochlea in these mice was similar to the cochlea of wild-type and *hOtx2¹/hOtx2¹* mice, the saccule was often smaller in size, indicating that *otd* might be less effective at compensating for *Otx1* functions than *Otx2* (Fig. 3E and Table 2).

Table 1. Genes affecting cochlear patterning

Gene	Type of protein	Distribution in the inner ear and surrounding structures	Mutant or knockout phenotype	Ref.
<i>Brn4</i>	Pou domain transcription factor	Periotic mesenchyme	Defects in fibroblasts of spiral ligament; shortened cochlea	48–51
<i>Dlx5</i>	Homeobox transcription factor	Dorsal region of otic vesicle; semicircular canals and endolymphatic duct	No anterior or posterior canal; reduced lateral canal; abnormal endolymphatic duct and cochlea	25, 26
<i>Eya 1</i>	Transcriptional coactivator	Ventralmedial otic vesicle; VIIIth ganglion; vestibular and cochlear sensory regions; periotic mesenchyme	No VIIIth ganglion; amorphic inner ear	18, 22, 52, 53
<i>Fgf3</i>	Growth factor	r5 and r6; prospective otic placode region; neurogenic and sensory regions	No endolymphatic duct or sac; reduced spiral ganglion; enlarged membranous labyrinth	47, 54–57
<i>Fgfr2 (IIIb)</i>	Growth factor receptor	Otic placode; dorsal and medial wall of otic vesicle; nonsensory regions of the inner ear	Dysgenesis of membranous labyrinth; rudimentary sensory patches and VIIIth ganglion; 50% of mutants lack endolymphatic duct	47
<i>Hoxa1</i>	Homeobox transcription factor	8 dpc: r3/4 boundary to spinal cord	No endolymphatic duct or sac; amorphic inner ear; no organ of Corti; reduced VIIIth ganglion	58–61
<i>Hoxa1/b1</i>	Homeobox transcription factors	<i>Hoxb1</i> : 8 dpc: r3/4 boundary to spinal cord; 9 dpc: expression up-regulated in r4	Amorphic inner ear; more severe phenotype than <i>Hoxa1</i> $-/-$ alone	62
<i>Hoxa2</i>	Homeobox transcription factor	r1/2 boundary to spinal cord; expression upregulated in r3 and r5	Membranous labyrinth appeared enlarged; scala vestibuli lacking or collapsed	63, 64
<i>Kreisler</i>	bZIP Transcription factor	r5 and r6	Misplaced otocyst; inner ear usually cyst-like; endolymphatic duct is often missing	56, 65–67
<i>ngn 1</i>	bHLH transcription factor	Anterolateral otic vesicle	No VIIIth ganglion; fusion of utricle and saccule; shortened cochlea	11, 40
<i>Otx1</i>	Transcription factor	Lateral wall of otic vesicle; lateral canal and ampulla; lateral wall of saccule and cochlea	No lateral canal or ampulla; no lateral crista; incomplete separation of utricle and saccule; misshapen saccule and cochlea	10, 29
<i>Otx2</i>	Transcription factor	Ventral tip of otic vesicle; lateral wall of saccule and cochlea	<i>Otx1</i> $-/-$, <i>Otx2</i> $+/-$: more severe saccular and cochlear phenotype than <i>Otx1</i> $-/-$	10
<i>Pax2</i>	Paired-box transcription factor	Medial wall of otic vesicle; endolymphatic duct and sac; cochlea	Agenesis of the cochlea and spiral ganglion	17, 19, 68
<i>Pax3</i>	Paired-box transcription factor	Dorsal half of neural tube	Spot mouse: aberrant endolymphatic duct; misshapen cochlear and vestibular components	69–73

When the human *Otx1* cDNA was introduced into a disrupted *Otx2* locus (*hOtx1²/hOtx1²*), embryogenesis proceeded much further than in *Otx2*-null mice (36). The expression of *Otx1* in the visceral endoderm was able to rescue gastrulation and specification of rostral neuroectoderm that were defective in *Otx2* $-/-$ mice. However, despite the presence of *Otx1* mRNA, no OTX1 protein was detected in the epiblast of these mice. As a result, *hOtx1²/hOtx1²* mice lacked forebrain and midbrain structures, displaying a headless phenotype from 9 dpc onward. Among all of the specimens examined between 15 to 16.5 dpc for inner ear defects, the coiling and the shape of the cochlea were invariably affected (Figs. 4 and 5). Most specimens show an incomplete separation of the utricle and saccule (Table 2 and Fig. 4). The shape of the saccule often appeared thinner than that of wild type (Fig. 4A), and sometimes displayed aberrant notches (Fig. 4B, arrowhead). Compared with saccules of *hOtx2¹/hOtx2¹* mice, they were also more affected (compare Fig. 3D with F and Fig. 4). The distribution of human OTX1 protein in the *hOtx1²/hOtx1²* inner ears has not been examined, so the inability of *Otx1* to substitute for *Otx2* functions in the inner ear could also be caused by a posttranscriptional problem, similar to the situation in the epiblast. The lateral canal and ampulla in these inner ears were normal; an expected result because the *Otx1* locus was not

disrupted in these mice (Fig. 4). Taken together, these results suggest that *Otx1* and *Otx2* have both overlapping and specific functions in the patterning of the inner ear. *Otx1* is essential for the formation of the lateral canal and ampulla, whereas *Otx2* plays a critical role in the patterning of ventral structures such as the cochlea and saccule (Table 2).

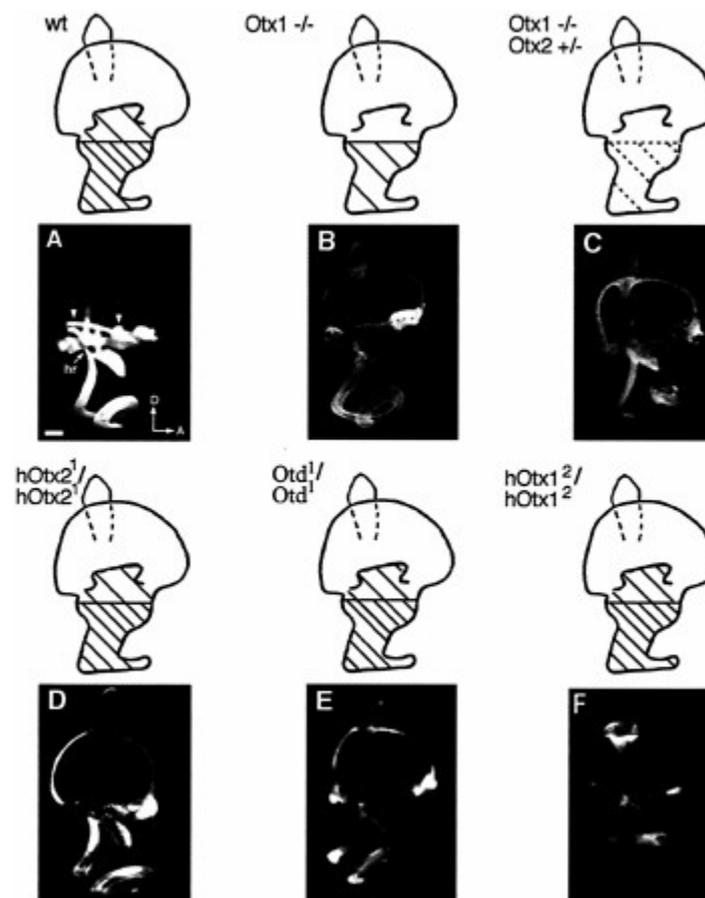


Fig. 3. Lateral views of paint-filled inner ears from wild-type (A), *Otx1*^{-/-} (B), *Otx1*^{-/-}; *Otx2*^{+/-} (C), *hOtx2*¹/*hOtx2*¹ (D), *otd*¹/*otd*¹ (E) and *hOtx1*²/*hOtx1*² (F) mice. Domains of *Otx1* (blue), *Otx2* (red), and *otd* (green) expression are shown as schematics above examples of paint-filled inner ears. In *Otx1*^{-/-} mutants (B), the lateral canal and ampulla are missing. The utricle and saccule are incompletely separated, and the shapes of the saccule and cochlea are often malformed. In *Otx1*^{-/-}; *Otx2*^{+/-} mutants (C), the phenotypes of the saccule and cochlea are more severe than those in *Otx1*^{-/-} mice. In *hOtx2*¹/*hOtx2*¹ mice (D), the lateral canal and ampulla are missing. The shapes of the saccule and cochlea are normal but the cochlea is sometimes shortened. In *otd*¹/*otd*¹ mice (E), the phenotype is similar to that of *hOtx2*¹/*hOtx2*¹, with no lateral canal and ampulla formation. In addition, the saccule also is malformed. In *hOtx1*²/*hOtx1*² mice (F), both the saccule and cochlea are malformed but the lateral canal and ampulla are normal, hr, hook region. Arrowheads in A indicate the lateral canal and ampulla, which are missing in B, C, D, and E. Orientation as per Fig. 1. (Scale bar=300 μm.)

Genes Expressed in the VIIIth Cranial Ganglion/Neurogenic Region. Neurons of the VIIIth ganglion are derived from otic epithelial cells that delaminate from the antero-lateral region of the otic cup/otocyst (37). Thus far, there is no direct evidence suggesting that normal formation of the VIIIth ganglion affects inner ear development. However, from gene expression studies, the neurogenic region and the presumptive sensory organs of the utricle, saccule, and cochlea most likely share a common *Lunatic fringe* (*L-fng*) expression domain (4, 38, 39). Interestingly, mice with a deletion of a basic helix-loop-helix gene, *neurogenin 1* (*ngn 1*), fail to form the VIIIth ganglion, and maculae of both the utricle and saccule are smaller in size (11). The length of the cochlear duct is also shorter compared with that of wild type. An attractive interpretation of these results is that the absence of *ngn 1* causes the loss of progenitor cells that normally give rise to sensory neurons as well as sensory hair cells and supporting cells

(11). As a result, defects are observed in both the ganglion and sensory epithelia. However, even though *ngn 1* is expressed in the expected neurogenic region in the otocyst stage (40), the expression of *ngn 1* in later stages of inner ear development has not been reported. Therefore, it is equally likely that *ngn 1* is expressed in both the neurogenic and sensory regions, and that this gene is independently required for the normal development of these regions. In this scenario, the development of the neurogenic and sensory fates are not related.

Table 2. Frequencies of various phenotypes in inner ears of *Otx1* and *Otx2* mutants

Genotype	No. of animals	Lack of lateral canal and ampulla	Lack of separation of utricle and saccule	Lack or aberrant cochleo-saccular duct	Aberrant saccule	Misshaoen cochlea		
						Hook region	No. of coils	Aberrant shape
<i>Otx1</i> ^{-/-}	11	11	11	11	6	11	9	5
<i>Otx1</i> ^{-/-}	9	9	9	9	7	9	8	7
<i>Otx2</i> ^{+/-} <i>hOtx2</i> ^{1/1}	7	7	6	5	0	5	1	0
<i>Otd</i> ^{1/Otd} ¹	6	6	6	5	6	5	0	0
<i>hOtx1</i> ^{2/2}	10	0	9	9	10	10	9	10

All mutant mice were scored between 1.5 to 16.5 dpc. At these stages, the cochlea should coil from 1.5 to a mature 1.75 turns. Any cochlea that had 1.5 turns was considered normal in order to accommodate for variability in staging and possible developmental delay of mutants.

In addition, the absence of the spiral ganglion may affect the proper formation of the modiolus (the bony tube that forms the central axis of the cochlea), which in turn may have a secondary effect on the final shape of the cochlea. In *ngn 1*-knockout mice, the modiolus is also missing and the coiling of the cochlea is tighter than is observed in wild-type mice (11). Despite the defects in the ganglion and sensory organs of *ngn 1*-mutant mice, the sensory hair cells appeared normal. This observation is consistent with the idea that normal innervation is not required for hair cell differentiation, at least until birth (41, 42).

Relationship Between Sensory Organ Specification and Gross Patterning of the Inner Ear. Two lines of evidence suggest that in the developing inner ear, sensory tissues are specified before nonsensory structures. First, there are no examples of either zebrafish or mouse mutants in which nonsensory structures develop normally in the absence of any sensory tissues. However, there are examples of mutants that lack nonsensory structures but develop normal sensory structures (12, 43–45). These observations suggest that nonsensory structures do not develop without the prior specification of some sensory tissues.

The second line of evidence stems from transplantation experiments. When the antero-posterior (A/P) axis of the chicken inner ear is surgically reversed at a stage when the otocyst is almost closed, the A/P axis of the sensory organs in this transplanted inner ear is already specified by the donor. In contrast, the A/P axis of the nonsensory structures such as the semicircular canals are respecified according to the new axial information from the host (46). As a result, in such a transplanted inner ear, the posterior crista, for example, is now located in the anterior region of the inner ear, and the posterior semicircular canal that is normally connected to the posterior ampulla is positioned anteriorly and adopts the pattern of an anterior canal. This evidence strongly suggests that there is a temporal delay in the specification of sensory versus nonsensory tissues.

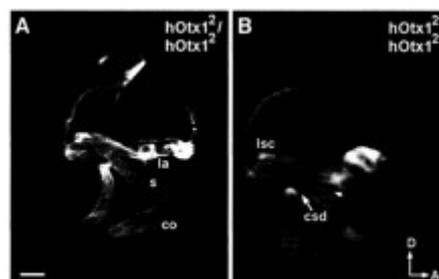


Fig. 4. Lateral views of paint-filled *hOtx1*^{2/2} inner ears from 16 (A) and 16.5 (B) dpc. The lateral canal and ampulla are normal. The saccule and cochleosaccular ducts are affected in both specimens, but B is more severe than A is. Arrowhead in B indicates an aberrant notch in the saccule. Orientation and abbreviations as per Fig. 1. (Scale bar=200 μ m.)

Thus far, there are no reported mutants in zebrafish or mice that have normal semicircular canals (nonsensory) but lack their corresponding sensory tissues, the cristae (12, 43–45). This observation raises an interesting possibility that the development of nonsensory tissues within the inner ear is under the influence of sensory structures. The identification of such signaling molecules produced by sensory tissues will be essential in unraveling the formation of this complex organ. Recently, it has been suggested that fibroblast growth factor (FGF) 10, produced in the sensory regions, is one of the ligands directing development of adjacent nonsensory structures that express its receptor, *Fgfr2* (47). These results are supported by the fact that knockout of *Fgfr2 IIIb*, one of the two functional isoforms of *Fgfr2*, yielded an inner ear with poor vestibular as well as cochlear development. However, sensory patches in these mutant mice are also rudimentary, suggesting that nonsensory tissues may also feedback on sensory tissues for their further development. As more and more of these signaling molecules are identified, it should be feasible to establish a hierarchy of molecular events starting from otic induction to a mature inner ear.

Most existing inner ear mutants display defects in both sensory and nonsensory structures (12, 43–45). The genes involved may play a role in specifying or coordinating sensory and nonsensory development. Depending on the domain of expression and the type of gene product, these genes could also be independently required for the formation of the structures involved. In *Otx1*-knockout mice, both the lateral canal and crista do not form (10). Gene expression and paint-fill data suggest that the presumptive lateral crista and the lateral canal are present initially but fail to develop in the mutant. Therefore, *Otx1* is most likely playing a role in the continued development of the prespecified lateral crista and canal. It remains to be determined whether *Otx1* plays a role in coordinating the development of these two structures (10).

Conclusion

Knockout and mutant mice will continue to be an indispensable tool in understanding normal development of the inner ear. However, to decipher the molecular mechanisms that underlie the normal developmental process, efforts must be invested beyond

mere documentation of mutant phenotypes. For any given gene, it is important to determine where along the developmental cascade a given gene acts, with whom it interacts, and how its functions. Correlating pattern of expression with phenotype is a first step toward achieving that goal. More sophisticated gene targeting approaches designed to remove gene functions in a spatially or temporally restricted manner will also facilitate the deciphering of the development of this complex organ.

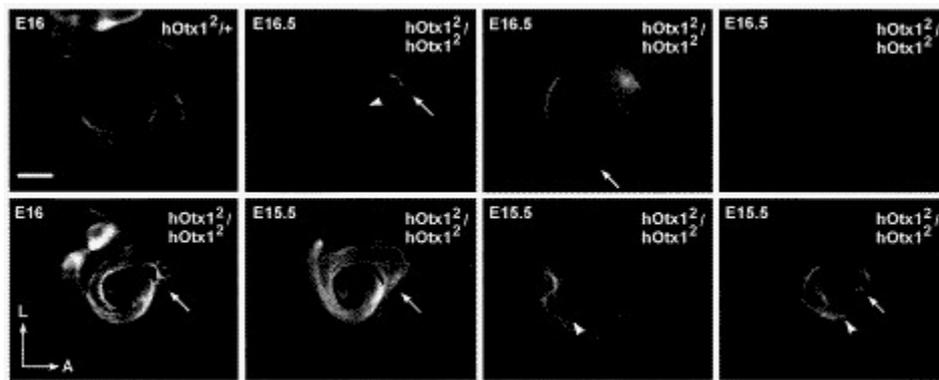


Fig. 5. Ventral views of cochleae from heterozygous and *hOtx1²/hOtx1²* mutant mice. Each of the mutant cochleae demonstrates abnormalities in both number of coils and shape (arrows). In addition, some coils have aberrant protrusions (arrowheads). *En*, embryonic day n. Orientation: A, anterior; L, lateral. (Scale bar=200 μ m.)

We thank Drs. Susan Sullivan and Bernd Fritzsich for critically reviewing this manuscript.

- Hudspeth, A.J. (1989) *Nature (London)* **341**, 397–404.
- Echteler, S.M., Fay, R.R. & Popper, A.N. (1994) in *Comparative Hearing: Mammals*, eds. Fay, R.R. & Popper, A.N. (Springer, New York), pp. 134–171.
- Lewis, E.R., Leverenz, E.L. & Bialek, W. (1985) *The Vertebrate Inner Ear* (CRC, Boca Raton, FL).
- Morsli, H., Choo, D., Ryan, A., Johnson, R. & Wu, D.K. (1998) *J.Neurosci.* **18**, 3327–3335.
- Ruben, R.J. (1967) *Acta Otolaryngol. Suppl.* **220**, 1–44.
- Wangeman, P. & Schacht, J. (1996) in *The Cochlea*, eds. Dallos, P., Popper, A.N. & Fay, R.R. (Springer, New York), Vol. 8, p. 551.
- Cowan, C.A., Yokoyama, N., Bianchi, L.M., Henkemeyer, M. & Fritzsich, B. (2000) *Neuron* **26**, 417–430.
- Hadrys, T., Braun, T., Rinkwitz-Brandt, S., Arnold, H.H. & Bober, E. (1998) *Development (Cambridge, U.K.)* **125**, 33–39.
- Wang, W., Van De Water, T. & Lufkin, T. (1998) *Development (Cambridge, U.K.)* **125**, 621–634.
- Morsli, H., Tuorto, F., Choo, D., Postiglione, M.P., Simeone, A. & Wu, D.K. (1999) *Development (Cambridge, U.K.)* **126**, 2335–2343.
- Ma, Q., Anderson, D.J. & Fritzsich, B. (2000) *J.Assoc. Res. Otolaryngol.*, in press.
- Fekete, D.M. (1999) *Trends Neurosci.* **22**, 263–269.
- Torres, M. & Giraldez, F. (1998) *Mech. Dev.* **71**, 5–21.
- Fritzsich, B., Barald, K.F. & Lomax, M.I. (1998) in *Springer Handbook of Auditory Research*, eds. Rubel, E.W., Popper, A.N. & Fay, R.R. (Springer, New York), Vol. 12, pp. 81–145.
- Fekete, D.M. (1996) *Curr. Opin. Neurobiol.* **6**, 533–541.
- Noden, D.M. & Van de Water, T.R. (1992) *Trends Neurosci.* **15**, 235–237.
- Torres, M., Gomez-Pardo, E. & Gruss, P. (1996) *Development (Cambridge, U.K.)* **122**, 3381–3391.
- Xu, P.X., Adams, J., Peters, H., Brown, M.C., Heaney, S. & Maas, R. (1999) *Nat. Genet.* **23**, 113–117.
- Rinkwitz-Brandt, S., Arnold, H.H. & Bober, E. (1996) *Hear. Res.* **99**, 129–138.
- Nornes, H.O., Dressler, G.R., Knapik, E.W., Deutsch, U. & Gruss, P. (1990) *Development (Cambridge, U.K.)* **109**, 797–809.
- Mansouri, A., Chowdhury, K. & Gruss, P. (1998) *Nat. Genet.* **19**, 87–90.
- Xu, P.X., Woo, I., Her, H., Beier, D.R. & Maas, R.L. (1997) *Development (Cambridge, U.K.)* **124**, 219–231.
- Simeone, A., Acampora, D., Pannese, M., D'Esposito, M., Stornaiuolo, A., Gulisano, M., Mallamaci, A., Kastury, K., Druck, T., Huebner, K., et al. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2250–2254.
- Robinson, G.W. & Mahon, K.A. (1994) *Mech. Dev.* **48**, 199–215.
- Acampora, D., Merlo, G.R., Paleari, L., Zerega, B., Postiglione, M.P., Mantero, S., Bober, E., Barbieri, O., Simeone, A. & Levi, G. (1999) *Development (Cambridge, U.K.)* **126**, 3795–3809.
- Depew, M.J., Liu, J.K., Long, J.E., Presley, R., Meneses, J.J., Pedersen, R.A. & Rubenstein, J.L. (1999) *Development (Cambridge, U.K.)* **126**, 3831–3846.
- Qiu, M., Bulfone, A., Martinez, S., Meneses, J.J., Shimamura, K., Pedersen, R.A. & Rubenstein, J.L. (1995) *Genes Dev.* **9**, 2523–2538.
- Morasso, M. I., Grinberg, A., Robinson, G., Sargent, T.D. & Mahon, K.A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 162–167.
- Acampora, D., Mazan, S., Avantiaggiato, V., Barone, P., Tuorto, F., Lallemand, Y., Brulet, P. & Simeone, A. (1996) *Nat. Genet.* **14**, 218–222.
- Acampora, D., Mazan, S., Lallemand, Y., Avantiaggiato, V., Maury, M., Simeone, A. & Brulet, P. (1995) *Development (Cambridge, U.K.)* **121**, 3279–3290.
- Royet, J. & Finkelstein, R. (1995) *Development (Cambridge, U.K.)* **121**, 3561–3572.
- Hirth, F., Therianos, S., Loop, T., Gehring, W.J., Reichert, H. & Furukubo-Tokunaga, K. (1995) *Neuron* **15**, 769–778.
- Ang, S.L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. & Rossant, J. (1996) *Development (Cambridge, U.K.)* **122**, 243–252.
- Acampora, D., Avantiaggiato, V., Tuorto, F., Barone, P., Perera, M., Choo, D., Wu, D., Corte, G. & Simeone, A. (1999) *Development (Cambridge, U.K.)* **126**, 1417–1426.
- Acampora, D., Avantiaggiato, V., Tuorto, F., Barone, P., Reichert, H., Finkelstein, R. & Simeone, A. (1998) *Development (Cambridge, U.K.)* **125**, 1691–1702.
- Acampora, D., Avantiaggiato, V., Tuorto, F., Briata, P., Corte, G. & Simeone, A. (1998) *Development (Cambridge, U.K.)* **125**, 5091–5104.
- Carney, P.R. & Couve, E. (1989) *Anat. Rec.* **225**, 156–164.
- Cole, L. K., Le Roux, L., Nunes, F., Laufer, E., Lewis, J. & Wu, D.K. (2000) *J.Comp. Neurol* **42**, 509–520.
- Adam, J., Myat, A., Le Roux, I., Eddison, M., Henrique, D., Ish-Horowicz, D. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 4645–4654.
- Ma, Q., Chen, Z., Barrantes, L., de la Pompa, J.L. & Anderson, D.J. (1998) *Neuron* **20**, 469–482.
- Fritzsich, B., Silos-Santiago, I., Bianchi, L.M. & Farinas, I. (1997) *Semin. Cell Dev. Biol.* **8**, 277–284.
- Silos-Santiago, I., Fagan, A.M., Garber, M., Fritzsich, B. & Barbacid, M. (1997) *Eur. J.Neurosci.* **9**, 2045–2056.
- Whitfield, T.T., Granato, M., van Eeden, F.J., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C.P., Jiang, Y.J., et al. (1996) *Development (Cambridge, U.K.)* **123**, 241–254.

44. Malicki, J., Schier, A.F., Solnica-Krezel, L., Stemple, D.L., Neuhauss, S.C., Stainier, D.Y., Abdelilah, S., Rangini, Z., Zwartkruis, F. & Driever, W. (1996) *Development (Cambridge, U.K.)* **123**, 275–283.
45. Deol, M.S. (1983) in *Development of Auditory and Vestibular Systems*, ed. Romand, R. (Academic, London), pp. 309–333.
46. Wu, D.K., Nunes, F.D. & Choo, D. (1998) *Development (Cambridge, U.K.)* **125**, 11–20.
47. Pirvola, U., Spencer-Dene, B., Xing-Qun, L., Kettunen, P., Thesleff, L., Fritzsche, B., Dickson, C. & Ylikoski, J. (2000) *J. Neurosci.* **20**, 6125–6134.
48. de Kok, Y.J., van der Maarel, S.M., Bitner-Glindzicz, M., Huber, I., Monaco, A.P., Malcolm, S., Pembrey, M.E., Ropers, H.H. & Cremers, F.P. (1995) *Science* **267**, 685–688.
49. Phippard, D., Heydemann, A., Lechner, M., Lu, L., Lee, D., Kyin, T. & Crenshaw, E. B. 3rd (1998) *Hear. Res.* **120**, 77–85.
50. Phippard, D., Lu, L., Lee, D., Saunders, J.C. & Crenshaw, E.B., 3rd (1999) *J. Neurosci.* **19**, 5980–5989.
51. Minowa, O., Ikeda, K., Sugitani, Y., Oshima, T., Nakai, S., Katori, Y., Suzuki, M., Furukawa, M., Kawase, T., Zheng, Y., *et al.* (1999) *Science* **285**, 1408–1411.
52. Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., Weil, D., Cruaud, C., Sahly, I., Leibovici, M., *et al.* (1997) *Nat. Genet.* **15**, 157–164.
53. Kalatzis, V., Sahly, I., El-Amraoui, A. & Petit, C. (1998) *Dev. Dyn.* **213**, 486–499.
54. Mansour, S.L. (1994) *Mol. Reprod Dev.* **39**, 62–68.
55. Mansour, S.L., Goddard, J.M. & Capecchi, M.R. (1993) *Development (Cambridge, U.K.)* **117**, 13–28.
56. McKay, I.J., Lewis, J. & Lumsden, A. (1996) *Dev. Biol.* **174**, 370–378.
57. Wilkinson, D.G., Bhatt, S. & McMahon, A.P. (1989) *Development (Cambridge, U.K.)* **105**, 131–136.
58. Mark, M., Lufkin, T., Vonesch, J.L., Ruberte, E., Olivo, J.C., Dolle, P., Gorry, P., Lumsden, A. & Chambon, P. (1993) *Development (Cambridge, U.K.)* **119**, 319–338.
59. Murphy, P. & Hill, R.E. (1991) *Development (Cambridge, U.K.)* **111**, 61–74.
60. Chisaka, O., Musci, T.S. & Capecchi, M.R. (1992) *Nature (London)* **355**, 516–520.
61. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. (1991) *Cell* **66**, 1105–1119.
62. Gavalas, A., Studer, M., Lumsden, A., Rijli, F.M., Krumlauf, R. & Chambon, P. (1998) *Development (Cambridge, U.K.)* **125**, 1123–1136.
63. Maconochie, M., Nonchev, S., Morrison, A. & Krumlauf, R. (1996) *Annu. Rev. Genet.* **30**, 529–556.
64. Rijli, F.M., Mark, M., Lakkaraju, S., Dierich, A., Dolle, P. & Chambon, P. (1993) *Cell* **75**, 1333–1349.
65. Cordes, S.P. & Barsh, G.S. (1994) *Cell* **79**, 1025–1034.
66. Deol, M.S. (1964) *J. Embryol. Exp. Morph.* **12**, 475–490.
67. Ruben, R. (1973) *Laryngoscope* **83**, 1440–1468.
68. Favor, J., Sandulache, R., Neuhauser-Klaus, A., Pretsch, W., Chatterjee, B., Senft, E., Wurst, W., Blanquet, V., Grimes, P., Sporle, R. & Schughart, K. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13870–13875.
69. Goulding, M.D., Chalepakis, G., Deutsch, U., Erselius, J.R. & Gruss, P. (1991) *EMBO J.* **10**, 1135–1147.
70. Goulding, M.D., Lumsden, A. & Gruss, P. (1993) *Development (Cambridge, U.K.)* **117**, 1001–1016.
71. Goulding, M., Sterrer, S., Fleming, J., Balling, R., Nadeau, J., Moore, K.J., Brown, S.D., Steel, K.P. & Gruss, P. (1993) *Genomics* **17**, 355–363.
72. Tassabehji, M., Read, A.P., Newton, V.E., Harris, R., Balling, R., Gruss, P. & Strachan, T. (1992) *Nature (London)* **355**, 635–636.
73. Deol, M. (1966) *Nature (London)* **209**, 219–220.

CELLULAR STUDIES OF AUDITORY HAIR CELL REGENERATION IN BIRDS

Jennifer S. Stone and Edwin W. Rubel*

Virginia Merrill Bloedel Hearing Research Center and Department of Otolaryngology, University of Washington School of Medicine, Seattle, WA 98195-7923

A decade ago it was discovered that mature birds are able to regenerate hair cells, the receptors for auditory perception. This surprising finding generated hope in the field of auditory neuroscience that new hair cells someday may be coaxed to form in another class of warm-blooded vertebrates, mammals. We have made considerable progress toward understanding some cellular and molecular events that lead to hair cell regeneration in birds. This review discusses our current understanding of avian hair cell regeneration, with some comparisons to other vertebrate classes and other regenerative systems.

supporting cell | avian | basilar papilla

Hair Cells: Old and New

Hair cells are anatomically and functionally exquisite cells that serve as the mechanoreceptors for hearing, balance, and orientation in space. Their name is derived from the bundle of actin-filled stereocilia that protrudes from their apical surfaces into the fluid-filled cavities of the inner ear or lateral line organs. Each hair cell is surrounded by several supporting cells and forms synapses with the VIIIth cranial nerve, which encodes and transmits signals to and from the hindbrain nuclei. Auditory hair cells are highly susceptible to intense noise, ototoxic drugs, and aging, and many genetic defects lead to malformations of the peripheral auditory structures. The majority of hearing loss in humans is sensorineural in nature—i.e., derived from abnormalities in either hair cells or the VIIIth nerve. Unfortunately, hair cell loss in humans and most mammals is irreversible.

The first indication that warm-blooded vertebrates can regenerate hair cells in the inner ear came from relatively recent experiments aimed at addressing two completely different sets of problems in birds. Cruz *et al.* (1) were attempting to use the chicken cochlea (or basilar papilla) as a model system to study ototoxicity of aminoglycoside antibiotics. At the same time, Cotanche (2) was studying developmental changes in noise damage and tonotopic organization in the same tissue. Both groups reported preliminary indications that new hair cells were formed in the area of hair cell damage in the posthatch chick basilar papilla. Confirmation of this interpretation was provided when radioactive thymidine labeling was seen in both hair cells and supporting cells after noise damage in young chicks (3) and adult quail (4) (Fig. 1). At the same time, Jørgensen and Mathiesen (5) showed that there was mitotic activity and new hair cell production in the vestibular epithelia of untreated adult parakeets (Budgerigars). The vestibular epithelium's capacity for ongoing hair cell production is distinct from the avian auditory epithelium, in which there is no postembryonic mitotic activity (6) until hair cell damage is induced. Similar to basilar papillas, the rate of hair cell production is increased in avian vestibular organs after experimental damage (7–11).

Methods for Stimulating Hair Cell Regeneration in the Basilar Papilla

This paper focuses on hair cell regeneration in the avian auditory epithelium. Historically, two principal experimental methods have been used to induce the loss of auditory hair cells in birds: exposure to intense noise and treatment with ototoxic drugs.

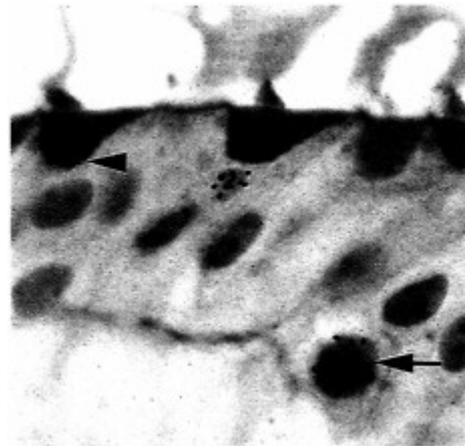


Fig. 1. Evidence of mitotic regeneration of hair cells and supporting cells in the avian basilar papilla. The nuclei of hair cells (arrowhead) and supporting cells (arrow) contain the radioactive nucleotide analog, tritiated thymidine, after noise damage in the mature quail, demonstrating they are the products of renewed cell division.

Exposure to intense pure-tone or broadband noise causes a lesion that varies with frequency along the tonotopic axis (length) of the epithelium, and that has a size and location along the neural axis (width) of the epithelium that change with intensity. The lesions that result from sound exposures at high intensities are composed of regions of complete hair cell loss as well as regions in which hair cells survive but appear damaged in a variety of ways (12–14). In most cases, noise exposure affects additional cell types besides hair cells (15–17).

The second method of inducing hair cell loss uses the ototoxic antibiotics, aminoglycosides (e.g., gentamicin, kanamycin, and streptomycin). Aminoglycosides target hair cells in the high-frequency, proximal portion of the basilar papilla. At low single doses, only hair cells at the proximal tip of the basilar papilla are

*To whom reprint requests should be addressed at: Virginia Merrill Bloedel Hearing Research Center, CHDD CD 176, Box 357923, University of Washington, Seattle, WA 98195-7923. E-mail: rubel@u.washington.edu.

This paper was distributed at and is part of the papers presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: FGF, fibroblast growth factor; IGF, insulin-like growth factor; FGFR, FGF receptor; RA, retinoic acid.

killed (18). As dosage is increased, the region devoid of native hair cells expands toward the distal end. Recently, several laboratories have developed and standardized single-dosage, systemic, or local application paradigms for aminoglycosides (18–22). By carefully documenting the region of complete hair cell loss with each drug administration paradigm, it is possible to ensure that all hair cells in a particular frequency region of the basilar papilla are regenerated. Aminoglycosides generally cause complete hair cell loss throughout most of the lesioned area, with intermittent hair cell loss at the distal border of the lesion. At moderate doses they appear to induce little or no direct damage to nonhair cell elements of the avian sensory epithelium.

Avian Hair Cell Progenitors: Identity and Behavior

The observation that new hair cells and supporting cells are formed after noise exposure led to the hypothesis that supporting cells are the progenitors to new hair cells during avian regeneration (2–4). Direct evidence was derived from studying the earliest cell types to enter S phase after experimentally induced damage in chickens. After noise damage, both supporting cells and hyaline cells, which line the abneural edge of the auditory epithelium, incorporate exogenous nucleotides during early phases of regeneration (15, 23, 24). Subsequent analysis has ruled out hyaline cells as potential hair cell progenitors, because hyaline cell proliferation and migration into the region of hair cell extrusion after intense noise stimulation do not lead to the formation of new hair cells (25). Observations of the chick utricle, a vestibular organ (26), and the basilar papilla after laser ablation (27) or ototoxic drug treatment (20, 28) also revealed that supporting cells are the principal epithelial cells to enter S phase preceding hair cell regeneration.

Supporting cell entry into the cell cycle is detected by 18–24 h after the onset of the damaging stimulus to the basilar papilla (15, 23, 24, 27, 29, 30). This delay reflects the latency of the triggering stimulus for cell proliferation plus the time required for supporting cells to progress from the growth-arrested state to S phase. The level of supporting cell proliferation peaks within 2–3 days after the onset of the stimulus, and it returns to normal levels by 7–9 days (19, 23, 24, 30). Patterns of cell proliferation spatially and temporally mirror the progression of hair cell loss (31).

Although supporting cells appear to be the likely candidates for avian hair cell progenitors, it is not clear whether all supporting cells have this potential. After drug treatment, virtually all of the supporting cells in both the damaged and undamaged regions of the basilar papilla change their cell cycle status by leaving growth arrest and entering G₁ phase, as demonstrated by increased expression of proliferating cell nuclear antigen (19). This observation suggests that all supporting cells in some nonmammalian species have the potential to divide. Some support for this notion is provided by Presson *et al.* (32) in their study of the fish sacculle. They showed that new supporting cells are recruited to divide during ongoing hair cell production after cytosine arabinoside (AraC) treatment, which kills progenitor cells in S phase (e.g., ref. 33). This finding demonstrated that the full component of progenitor cells was not actively dividing at the time AraC was administered and suggested that there is extended potential for proliferation among the quiescent supporting cell population.

Despite these observations, only progenitors in areas of the chick basilar papilla with obviously damaged hair cells progress as far as S phase and regenerate new hair cells (2–4, 19, 27, 34). Further, the cells that do reach S phase comprise only about 15% of the supporting cell population in those areas (35). These findings raise two questions. First, is the supporting cell population subdivided into cells with different proliferative potentials (i.e., terminally differentiated cells, stem cells, or committed precursor cells)? Second, what mechanisms trigger the transitions of supporting cells from growth arrest to S phase?

There have been few systematic studies addressing the proliferative potential of supporting cells in the avian auditory epithelium, and the results of these studies are equivocal. Results from one study (24) suggest that some progenitor cells act like stem cells, undergoing multiple rounds of cell division after noise damage. A single injection of the thymidine analog, BrdUrd, was administered to chicks early during the regenerative process, and animals were allowed to recover for varying periods after the injection. The total number of BrdUrd-labeled cells increased significantly, and clusters of labeled cells appeared to grow in size, suggesting growth of a colony of cells from one progenitor. However, a second study (30) suggests that progenitor cell recycling is in fact rare in the drug-damaged basilar papilla. We detected only very limited incorporation of two nucleotide analogs into the same progenitor cells when each nucleotide was administered separately at intervals estimated to mimic the average cycling time of mature eukaryotic cells. In contrast, recycling of progenitor cells occurs to a very high degree (70%) in the sacculles of normal oscars (36) and to a more limited extent in the utricles of control chicks (30, 36).

Studies of the mammalian olfactory epithelium, which regenerates olfactory neurons on an ongoing basis and in response to damage, have shed considerable light on the profiles of neuronal progenitors in that tissue. Molecular analyses and studies of proliferative behavior of cultured olfactory epithelial cells have shown that at least three subtypes of progenitors exist among the supporting cells (reviewed in ref. 37). The neuronal colony-forming cells, which compose a small percentage of the progenitor pool (<0.1%), resemble true stem cells in that they divide continually at a slow rate for a long period (38). Two additional cell populations with more limited proliferative potential exist: a Mash1-positive cell that is a committed amplifying cell (39) and its progeny cell, termed an intermediate neuronal precursor, which gives rise directly to olfactory neurons (40–42). Our field would benefit from analyses like these in the olfactory epithelium to elucidate progenitor subtypes. In the absence of this clarity, the term “progenitor cell” will be used in place of the term “supporting cell” throughout this paper to describe the cell that divides and gives rise to new cells in the damaged auditory epithelium.

Several laboratories have participated in the search for molecules that trigger progenitor cell proliferation in avian inner ear epithelia. This search has been facilitated by the development of culturing methods for auditory and vestibular end organs and for isolated sensory epithelia (8, 20, 43–46). Tsue *et al.* (47) found increased incorporation of radioactive thymidine into progenitor cells in single control utricles that were cocultured without contact with multiple utricles in which hair cell damage had been stimulated. Further, focal laser ablation of a few hair cells *in vitro* causes activation of progenitor cell proliferation up to 100 μ m from the lesion site (27). These studies suggest that a diffusible stimulatory signal is released from inner ear epithelia at the site of hair cell injury.

The identity and source of the putative diffusible mitogen(s) responsible for these effects are not known. Supporting cells in cultured utricles continue to divide in the absence of serum (43, 44), suggesting that substances that are mitogenic for hair cell progenitors are intrinsic to the sensory organs. The VIIIth nerve synapses/fibers are damaged as a result of hair cell loss (48–50) and therefore seem to be good candidates for release of mitogenic substances. Such an effect appears to occur in other regions of the nervous system. For example, mitogens released from injured peripheral nerve processes are thought to induce Schwann cell proliferation (51). However, neural elements are not required for proliferation during hair cell regeneration, as mitotic activity proceeds in tissue cultures of auditory (20) and

vestibular (43, 52) epithelia isolated without the nerve. In addition, neural elements do not appear to play a critical role in the genesis of sensory epithelial cells. Transplantation of the otocyst without its connected ganglia does not appear to disrupt production of either hair cells or supporting cells (53, 54). In mice that are null for both brain-derived neurotrophic factor and neurotrophin-3, both hair cells and supporting cells form despite the complete absence of cochlear ganglion cells (55). It should be noted, however, that innervation may be required for normal differentiation of hair cells (e.g., see refs. 55–57) and for maintenance of hair cells once they are fully differentiated (e.g., see ref. 58).

What are potential mitogens for progenitor cells in the chick inner ear? Studies using reverse transcription-PCR, *in situ* hybridization, or immunolabeling have revealed that epithelial cells in the mature avian basilar papilla express the following potential diffusible mitogens: fibroblast growth factor (FGF)-1 (59), FGF-2 (59–61), and insulin-like growth factor (IGF)-1 (59). FGF-2 protein is present in supporting cell nuclei (61), and its levels appear to be highly up-regulated in the damaged area after noise exposure (60). These growth factors bind to and activate tyrosine kinase receptors, many of which also are expressed in the avian inner ear. PCR studies have shown that mRNAs for FGF receptor (FGFR)-1, epidermal growth factor receptor (erbB1), and IGF receptor-1 are present in the cochlear duct (61) and, more specifically, in the sensory epithelium (59). Further, *in situ* hybridization reveals abundant message for FGFR-3 in the supporting cells (unpublished observation). After hair cell damage, there are detectable changes in the levels of transcripts for FGF-1, IGF receptor-1, FGFR-1, and FGFR-3, but not for IGF-1, FGF-2, or ErbB1 (59, 61). At the protein level, FGFR-1 is present in hair cells in control tissue, and then it becomes elevated in supporting cells after hair cell damage (59, 61). However, the actual mitogenic properties of most of these growth factors in auditory hair cell regeneration have not yet been determined. A study from our laboratory has shown that addition of IGF-1 or insulin to cultured chick utricles causes progenitor cell proliferation to increase significantly (62), whereas it has no effect in the basilar papilla (E.Oesterle, personal communication).

The role of leukocytes as activators of progenitor cell proliferation also is being explored. In response to hair cell damage, there is substantial migration and proliferation of macrophages and microglia-like cells in the inner ear epithelia of chicks (11, 63) and other species (64–67). Leukocytes can produce growth factors and cytokines (68), which may act directly as mitogens or indirectly by stimulating growth factor production in target cells (e.g., ref. 69). Indirect evidence suggests that the secretory products of leukocytes also may stimulate mitotic activity during hair cell regeneration. Proliferation of leukocytes in the avian auditory and vestibular epithelia precedes the experimentally induced increase in progenitor cell proliferation (11). Further, tumor necrosis factor α , which is released by macrophages after tissue damage (68), induces an increase in progenitor cell proliferation when added to cultures of the chick utricular epithelium (70).

Binding of growth factors to extracellular receptors initiates a cascade of intracellular signals, leading to increased mitotic activity. One requirement for a full understanding of regulation of hair cell regeneration is to determine which signaling pathways downstream of the receptor are activated during progenitor cell proliferation. Exploration of these pathways is just beginning. Activation of cAMP leads to increased progenitor cell proliferation in the basilar papilla *in vitro* (45). Because cAMP activity has not been widely implicated in growth factor-mediated stimulation of cell proliferation in other systems, the manner in which increased cAMP signaling serves to stimulate mitosis in the chick sensory epithelium is not clear.

Recruitment of progenitor cells to the cell cycle also may occur in response to local changes in the cells' microenvironment, such as the alterations in direct cell/cell signaling and cell shape that occur during hair cell extrusion and the ensuing supporting cell expansion (71).

In addition to positive control of mitotic activity, antimitotic influences are likely to be at work in the avian auditory epithelium: (i) to prevent proliferation in the undamaged basilar papilla; (ii) to ensure that all supporting cells in the damaged area do not proliferate; and (iii) to down-regulate cell proliferation when the correct number of cells is regenerated. The fact that auditory hair cell regeneration requires the loss of hair cells has led to the hypothesis that healthy hair cells exert an inhibitory influence on the mitotic activity of progenitor cells that surround them (2). Accordingly, Belgian *Waterslager* canaries, whose basilar papillae sustain spontaneous low-level hair cell damage in all regions (72), display ongoing progenitor cell proliferation (73). As noted above, a low level of ongoing cell proliferation occurs in the normal chick vestibular organs (5, 7, 74). This activity is thought to occur in response to spontaneous continual loss of individual hair cells through programmed cell death (74).

Do putative mitotic inhibitors in the sensory epithelia of the avian inner ear act through diffusion or through direct cell-cell contact? Tsue *et al.* (47) found that progenitor cell proliferation is down-regulated in single drug-damaged chick utricles that are cocultured without contact with multiple control utricles, suggesting that an inhibitory molecule is secreted from control organs. The inhibitory factor thought to be active in that system has not been characterized. However, molecules with inhibitory potential have been identified. Addition of exogenous basic FGF-2 to cultures of control utricles and drug-damaged basilar papillae leads to significant down-regulation of cell proliferation in the sensory epithelium (75). As mentioned above, FGFR-1 (59, 61) and FGFR-3 (O.Bermingham-McDonogh, personal communication) are expressed by supporting cells in the basilar papilla, and both receptors can be activated by FGF-2. Interestingly, FGFR-3 expression is not seen in the avian vestibular epithelia, which displays continuous mitotic activity. Furthermore, FGFR-3 expression is down-regulated in the regenerating region of the basilar papilla after drug treatment (unpublished observations). These observations suggest that FGFR-3 may provide steady-state inhibitory influences over supporting cells in undamaged auditory tissues.

Contact-mediated signaling from hair cells to progenitor cells also may play a role in inhibiting proliferation, but there are currently no data to directly support this hypothesis. This phenomenon appears to occur in the regenerating rodent olfactory epithelium. Addition of differentiated olfactory neurons to cultures of olfactory neuronal progenitors induces decreased mitotic activity among the progenitor cells (38). The mechanism that guides this negative feedback of differentiated cells upon their progenitors has not been characterized. However, one molecule that may mediate this inhibition is the extracellular receptor, Notch.

Notch is an integral membrane receptor that binds and becomes activated by ligands on adjacent cells. Notch has been studied most extensively in the developing *Drosophila* nervous system, where it plays a role in lateral inhibition, as well as in cell lineage decisions and boundary formation (reviewed in ref. 76). There is emerging evidence that signaling through the Notch receptor is also critical for neurogenesis in vertebrates (reviewed in ref. 77). Recent studies have shown that Notch and some of its ligands are expressed in the basilar papilla during development (78) and regeneration (79), suggesting that cell-cell signaling via this receptor is important for hair cell genesis. Further, the developing sensory epithelia of mutant zebrafish (80, 81) and knockout mice (82) show abnormal cell fate specification and

tissue patterning, lending additional support for this hypothesis (see more in-depth discussion below).

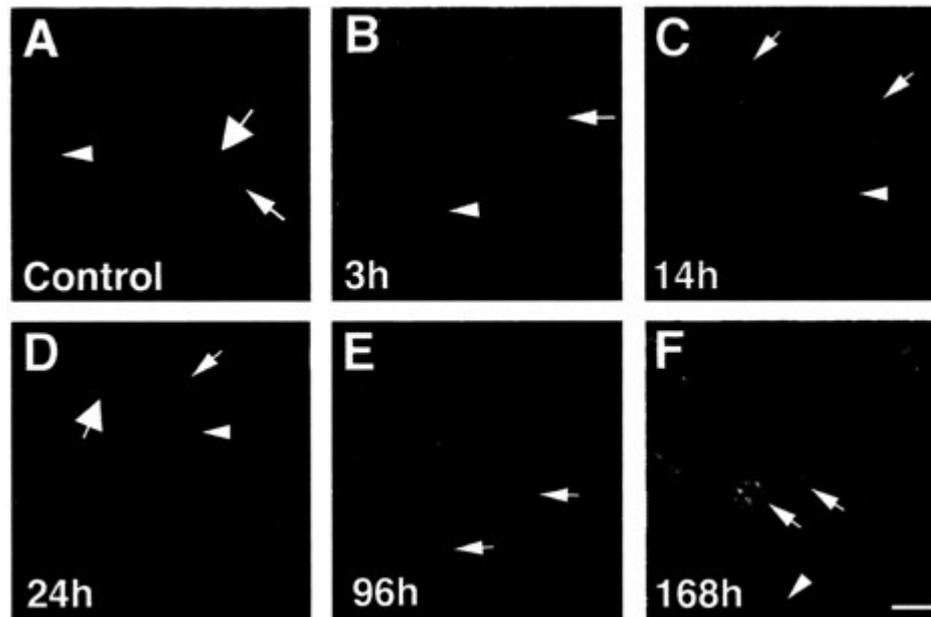


Fig. 2. Temporal progression of hair cell differentiation disclosed by antibodies to BrdUrd and the hair cell-selective protein, class —III β -tubulin. Whole-mount preparations of the basilar papilla (BP) labeled with antibodies to BrdUrd (red) and/or β -tubulin (green). (A) In the control BP, β -tubulin is present in hair cells (thin arrow) and nerves (arrowhead), but not supporting cells (thick arrow). (B–F) Drug-damaged BPs taken from chicks at different times after a single BrdUrd injection at 3 days postgentamicin. (B) 3 h postBrdU. Progenitor cells in S or G₂ phase of the cell cycle (arrow) are labeled with BrdUrd, but not β -tubulin. β -tubulin is present in nerves (arrowhead) remaining after hair cell loss. (C–F) BrdUrd labeling in regenerated cells at progressively later stages of differentiation. (C) 14 h post-BrdUrd. Some rounded cells (arrows) near the lumen are double-labeled and represent new hair cells at an early stage of differentiation. These cells are associated with nerve processes (arrowhead). (D) 24 h post-BrdUrd. Regenerated hair cells (thin arrow) are fusiform in shape and are associated with nerve processes (arrowhead). Some BrdUrd-positive cells are not labeled for β -tubulin (thick arrow); these cells are not differentiating as hair cells. (E) 96 h (4 days) post-BrdUrd. Regenerated hair cells have thick cytoplasmic processes that extend toward the lumen and basal lamina (arrows) of the epithelium. (F) 168 h (7 days) post-BrdUrd. Regenerated hair cells (arrows) resemble mature hair cells morphologically. Arrowhead points to nerve process. (Scale bar=10 μ m.)

A potential role of Notch in regulating mitotic activity in mature and developing tissues is beginning to emerge. Human gain-of-function mutations in *Notch* result in neoplasias in several tissues (reviewed in ref. 83), suggesting that Notch activation also may promote the proliferative state. Two recent studies report opposing effects of Notch on cell proliferation during formation of the *Drosophila* wing, depending on the stage of development. During early stages, activation of Notch stimulates cell proliferation (84), whereas during late development, activation of Notch leads to growth arrest (85). Clearly, functional studies of Notch in the avian inner ear are needed to test its potential role in regulating cell proliferation.

Repatterning the Mature Sensory Epithelium

The first newly formed hair cells appear in the regenerating avian basilar papilla by 3–4 days after the onset of a damaging stimulus (2, 15, 20, 28, 86) and continue to emerge in a manner that spatially and temporally mirrors hair cell loss (28, 86). Several weeks after damage the alternating array of supporting cells and hair cells is fairly well reestablished, with only minor deviations from the normal patterning, and hair cells appear fully differentiated (14, 15, 87, 88). At this time, the auditory nerve, which retracts after hair cell loss, has grown back to form synapses with the new hair cells (48, 50, 89–92). In the previous section, we discussed ways in which the appropriate number of new cells may be formed during hair cell regeneration in birds. This section addresses potential regulatory mechanisms for reestablishing the appropriate cell types and patterning in the avian auditory epithelium after damage.

A discussion of cellular differentiation during regeneration would not be complete without addressing the recent progress made toward identifying markers for mature and regenerating cells in the avian auditory epithelium. Cell-specific or cell-selective markers are critical for studies of tissue repair and regeneration for many reasons. They permit the analysis of early steps of hair cell differentiation, when mature morphological features of hair cells are not yet acquired. For example, a recent study from our lab using antibodies against the hair cell-selective protein, class III β -tubulin, has revealed that differentiation of regenerated hair cells resembles the embryonic process (93–95). New cells are mitotically generated at the luminal surface of the epithelium (96). Two to three days later they form extensions that reach toward the luminal and adluminal surfaces of the epithelium (Fig. 2). By 7 days after mitosis they attain their mature globular shape. Information gleaned from such studies with cell-selective markers can be used to generate and test hypotheses about cellular and molecular interactions guiding cell commitment and differentiation during hair cell regeneration.

Markers also facilitate the characterization of the region of hair cell loss, allowing one to determine whether the damaging stimulus induces complete hair cell extrusion from the epithelium or only partial hair cell injury. This caveat has proven to be important in hair cell regeneration research, because some experimental paradigms in mammals induce stereociliary damage and repair rather than hair cell loss and scarring (97, 98). Markers are also useful for studying specific cell types in culture (20, 52, 98–101), because considerable cellular dedifferentiation often occurs *in vitro*, which makes it difficult to identify cell types solely by morphological criteria.

Several markers that are specific for mature and differentiating hair cells in chicks have been characterized in studies of the developing and mature auditory epithelium (20, 28, 78, 95, 102–107). Some of these markers have been used in culture paradigms to study hair cell regeneration (20, 52). A few antibodies and cDNA probes for supporting cell-specific proteins in the basilar papilla also have been identified (20, 108–110). However, these markers have not been used widely in studies of hair cell regeneration. Rather, supporting cells have been identified based on the exclusion of labeling for hair cell-specific antigens (e.g., refs. 28, 52, 79, and 111). Based on this ambiguity, we will use the term “nonhair cell” to describe cells whose identity has been inferred by using this approach.

The observation that the nuclei of both hair cells and supporting cells contain radioactive thymidine several days after it is injected after noise exposure led to the hypothesis that a common progenitor exists for both cell types in the mature inner ear of chickens (3). Recent studies have confirmed this notion in the developing basilar papilla as well as in the regenerating basilar papilla and utricle. Fekete *et al.* (112) used retroviral lineage analysis in the chick otocyst to show that hair cells and supporting cells in the basilar papilla often arise from a single progenitor cell division. Additional studies have shown that similar asymmetric differentiation predominates as the mode of cell production in the normal chick utricle (7, 30, 79) and the salamander lateral line organ (67). Asymmetric cell production is used during the ontogeny of many tissue types in a wide variety of species (reviewed in ref. 113). One advantage of forming cells with distinct phenotypes during each mitotic event is that the development of cellular diversity is ensured. In addition, if one of the daughter cells is a progenitor cell, then asymmetric cell production also ensures that cells with proliferative potential are maintained throughout the course of development and, in some cases, into the postembryonic period.

Interestingly, modes of cell production are quite variable in the regenerating basilar papilla after drug treatment. Double-labeling for BrdUrd and a hair cell marker has shown that individual mitotic events at 3 days postdrug treatment are equally likely to result in any of the following combinations of cell types: one hair cell and one nonhair cell, two hair cells, or two nonhair cells (28). This trend changes at a later stage of regeneration. Among cells born at 5 days postdrug treatment, the majority (71%) differentiate into two nonhair cells (J.S.S., unpublished data). These observations suggest that, across sensory epithelia and over the course of regeneration after damage in the basilar papilla, there are distinct differences in the differentiative potential of progenitor cells and/or the extrinsic factors that influence the fates of the daughter cells. Recent studies have begun to identify both intrinsic and extrinsic factors that may be important for proper tissue patterning during hair cell regeneration.

Cell fate outcomes are influenced in part by signaling from the progeny cell's environment. As discussed above, lateral inhibition via the Notch signaling pathway is perhaps the best-understood mechanism of extrinsic regulation of cell fate determination among sensory cells. Studies of the developing inner ear epithelia implicate a role for Notch and its ligands, Delta and Jagged, in hair cell specification (reviewed in refs. 77 and 114–116). In embryonic fish, birds, and rodents, transcripts for Notch 1, 2, and 3 and the ligands, Delta1 and Serrate1, are expressed in the otocyst (78, 80, 82, 117–123). Loss of function of the *Delta* homologue, *DeltaA*, in zebrafish leads to the overproduction of hair cells at the expense of supporting cells (81). In another zebrafish mutant, *mindbomb*, Notch signaling also is reduced, and a similar phenotype to the *DeltaA* mutant results (80). Further, in the developing mouse organ of Corti, the ligand Jagged-2 (a Serrate homologue) appears to be critical for proper cell patterning, as mice that are null for the gene generate supernumerary hair cells (82).

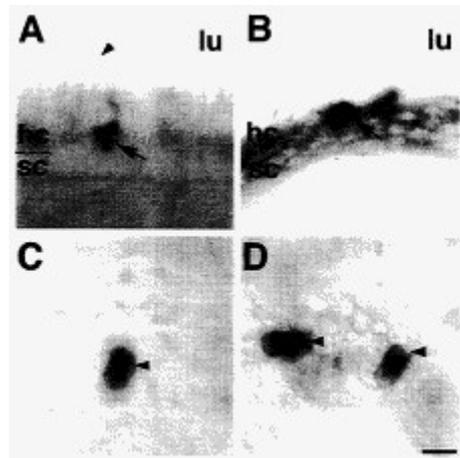


Fig. 3. Expression of *Delta1* mRNA in the chicken inner ear sensory epithelia. Cells expressing *Delta1* mRNA (black) are shown. Hair cells (arrows) in the untreated utricle (A) and drug-damaged basilar papilla at 3 days postgentamicin (B) contain high levels of *Delta1* mRNA. Arrowhead points to stereocilia at the luminal surface. Postmitotic sister pairs (arrowheads) in the untreated utricle (C) and drug-damaged basilar papilla at 3 days postgentamicin (D) have equal, high levels of *Delta1* mRNA. (Scale bar=5 μ m.) [Reproduced with permission from ref. 79 (Copyright 1999 The Company of Biologists Limited).]

Notch signaling in histogenesis in other adult tissues has been studied to only a limited degree (reviewed in refs. 83 and 124). In the regenerating inner ear epithelia of posthatch chickens, however, Notch signaling also may be critical for patterning the hair cells and supporting cells (79). Transcripts for Notch 1 and its ligand, Serrate1 (chick's Jagged 1 homologue), are abundant in nonregenerating inner ear tissues, whereas transcripts for *Delta1* are not. In mature chick utricles, and shortly after damage to the basilar papilla, *Delta1* mRNA is expressed in progenitor cells in S or G₂ phase, and it is symmetrically inherited by both daughter cells during mitosis (Fig. 3). During differentiation, *Delta1* mRNA levels appear to increase in cells acquiring the hair cell fate and disappear in cells acquiring the nonhair cell fate. During late stages of regeneration, *Delta1* mRNA expression decreases to its starting level. The effects of lateral inhibition may change over time in the regenerating basilar papilla and thereby influence the types of cells that are formed. For example, the increase in *Delta1*-expressing hair cells within the area of the lesion over time would lead to increased lateral inhibition and the generation of relatively more nonhair cells. Although the role of Delta1-Notch1 signaling in regenerating chick inner ear epithelia

remains to be tested experimentally, these observations suggest that the developmental role of Notch signaling appears to be preserved in the mature, regenerating chicken inner ear.

Signaling via diffusible factors is likely to be another important form of extrinsic regulation of cell specification during hair cell regeneration. One candidate factor, retinoic acid (RA), is a steroid hormone with long-range actions that is known to affect morphogenesis in many tissues (reviewed in ref. 125). RA, receptors for RA, and binding proteins for the precursor of RA, retinol, are present in the developing mammalian auditory epithelium (126–129). Addition of RA to cultures of the developing mouse organ of Corti causes an overproduction of hair cells (126), whereas inhibition of RA causes decreased hair cell specification (129). The role of RA in the organ of Corti of postnatal mammals remains controversial. LeFebvre *et al.* (130) reported that addition of RA to cultures of the drug-damaged cochlea of young rats stimulates regeneration or repair of hair cells, whereas Chardin and Romand (131, 132) detected no effect of RA in similar studies. In mature birds expression of RA receptors is present in the perinuclear region of hair cells (61), suggesting that RA signaling may be important for their differentiation. However, the role of RA in promoting cell specification in birds has not been tested.

The fate of newly regenerated cells depends on factors that are inherent to progenitor or progeny cells as well as extrinsic signaling molecules. These features include the activity of transcription factors and other cell fate determinants, which may be inherited during cell division (reviewed in refs. 113 and 133). Several transcription factors are expressed in the otocyst (reviewed in refs. 114, 134, and 135), but there is limited evidence that any of them is critical for cell fate determination or tissue patterning within the sensory epithelia. An exception is the helix-loop-helix transcription factor, *Math1*. Deletion of the *Math1* gene in mice leads to failure of hair cell production in both the auditory and vestibular epithelia, but supporting cells continue to be generated (136). Transfection of *Math1* in *in vitro* preparations of neonatal rat organ of Corti induces ectopic differentiation of hair cells in the region of the inner spiral sulcus (137). A role for *Math1* in avian hair cell regeneration remains to be determined. A recent series of experiments from our lab provides preliminary evidence that *cProx1*, a chicken homolog for the DNA binding protein *Prospero* in *Drosophila* (138), may be involved in hair cell production during development and regeneration (J.S.S., unpublished observations). *cProx1* protein is highly expressed in prosensory and proneural regions of auditory and vestibular portions of the avian otocyst, and expression is down-regulated by the time hair cell differentiation is complete. Interestingly, *cProx1* protein becomes highly and transiently re-elevated in the mature auditory epithelium after experimental damage, first in progenitor cells and later in hair cells. *Prospero* and its homologs have emerged as an interesting set of transcription factors with diverse roles. In the *Drosophila* nervous system, *Prospero* is necessary for proper cell specification and/or differentiation in some lineages (e.g., refs. 139 and 140). Future investigations should test the hypothesis that *cProx1* is required for hair cell specification and/or differentiation during regeneration.

Numb is a cytoplasmic protein whose activity within some cells is sufficient to confer a neural fate. Studies in *Drosophila* first showed that loss-of-function *Numb* mutants underproduce neurons (141, 142). Recently, homologues for *Numb* have been identified in other vertebrates (143–146), and a critical role for *Numb* in neural specification has been demonstrated in chickens (146). *Numb*'s effect on cell fate is determined at the time of cell division, as it is asymmetrically partitioned in the mother cell and distributed asymmetrically to daughter cells (reviewed in ref. 113). In recipient progeny, *Numb* acts by antagonizing Notch function (147–149). Based on the requirement for proper Notch signaling in hair cell specification, a potential role of *Numb* recently has begun to be explored in the developing and regenerating sensory epithelia of the inner ear. *Numb* protein is present in the mature chicken basilar papilla (J.S.S., unpublished observations). Further, *Numb* segregates asymmetrically within mitotic progenitor cells in this system. However, a critical role for *Numb* remains to be demonstrated.

Other cellular mechanisms, such as cellular reorganization (150) and programmed cell death (151), occur during development of the basilar papilla and may influence the numbers, types, or patterns of cells that are regenerated there in the mature bird after damage. Additionally, there is emerging evidence that hair cells may be produced by a nonmitotic conversion from supporting cells in mature frogs (152) and chickens (35, 153). This phenomenon, which is called direct or nonmitotic transdifferentiation, is addressed by Baird (161).

What's Wrong with Us?

Avian and mammalian auditory epithelia share many anatomical and functional features; they have similar types of specialized cell types, patterns of innervation, and mechanisms of sensory signal transduction. Despite these commonalities, these classes of vertebrates clearly possess critical differences with respect to their potential to form new cells after birth. The challenge lies in identifying where these differences exist. The most likely origin of divergence is the hair cell progenitor. In mammals, all cells in the auditory organ of Corti become terminally mitotic by embryonic day 14 (154). A similar phenomenon occurs in the avian basilar papilla by embryonic day 9 (155). Thus, the mature auditory organs in both classes of vertebrates exhibit mitotic quiescence. Despite this similarity, cells in the avian cochlear epithelium re-enter the cell cycle, divide, and differentiate into new hair cells after experimentally induced hair cell loss, but a similar stimulus causes no renewed mitotic activity in the mammalian organ of Corti (66). The failure of renewed proliferation in the mature mammalian auditory epithelium may be caused by persistent inhibition of mitotic activity among progenitor cells (156, 157), the absence of promitotic stimuli in response to hair cell loss (158–160), or the depletion of the hair cell progenitor during embryogenesis. The future's challenge is to identify which of these factors are the critical obstacles within the mammalian inner ear and to identify ways to overcome them. A growing understanding of the strategies used by animals with the capacity to regenerate inner ear hair cells, such as birds, will continue to provide insights into how this can be accomplished.

1. Cruz, R.M., Lambert, P.R. & Rubel, E.W (1987) *Arch. Otolaryngol. Head Neck Surg.* **113**, 1058–1062.
2. Cotanche, D.A. (1987) *Hear. Res.* **30**, 197–206.
3. Corwin, J.T. & Cotanche, D.A. (1988) *Science* **240**, 1772–1774.
4. Ryals, B.M. & Rubel, E. W (1988) *Science* **240**, 1774–1776.
5. Jørgensen, J.M. & Mathiesen, C. (1988) *Naturwissenschaften* **75**, 319–320.
6. Oesterle, E.C. & Rubel, E. W (1993) *Hear. Res.* **66**, 213–224.
7. Roberson, D.F., Weisleder, P., Bohrer, P.S. & Rubel, E. W (1992) *Hear. Res.* **57**, 166–174.
8. Oesterle, E.C., Tsue, T.T., Reh, T.A. & Rubel, E. W (1993) *Hear. Res.* **70**, 85–108.
9. Weisleder, P. & Rubel, E. W (1993) *J.Comp. Neurol.* **331**, 97–110.
10. Weisleder, P., Tsue, T.T. & Rubel, E. W (1995) *Hear. Res.* **82**, 125–133.
11. Bhave, S.A., Oesterle, E.C. & Coltrera, M.D. (1998) *J.Comp. Neurol.* **398**, 241–256.
12. Rubel, E. W & Ryals, B.M. (1982) *Acta Otolaryngol* **93**, 31–41.
13. Ryals, B.M. & Rubel, E. W (1982) *Acta Otolaryngol.* **93**, 205–210.
14. Cotanche, D.A., Petrell, A. & Picard, D.A. (1991) *Ciba Found. Symp.* **160**, 131–142.
15. Girod, D.A., Duckert, L.G. & Rubel, E. W (1989) *Hear. Res.* **42**, 175–194.
16. Poje, C.P., Sewell, D.A. & Saunders, J.C. (1995) *Hear. Res.* **82**, 197–204.
17. Cotanche, D.A. (1987) *Hear. Res.* **30**, 181–196.

18. Janas, J.D., Cotanche, D.A. & Rubel, E. W (1995) *Hear. Res* **92**, 17–29.
19. Bhave, S.A., Stone, J.S., Rubel, E. W & Coltrera, M.D. (1995) *J. Neurosci.* **15**, 4618–4628.
20. Stone, J.S., Leano, S.G., Baker, L.P. & Rubel, E. W (1996) *J. Neurosci.* **16**, 6157–6174.
21. Husmann, K.R., Morgan, A.S., Girod, D.A. & Durham, D. (1998) *Hear. Res.* **125**, 109–119.
22. Muller, M. & Smolders, J.W. (1999) *Hear. Res.* **131**, 153–169.
23. Hashino, E. & Salvi, R. (1993) *J. Cell Sci.* **105**, 23–31.
24. Stone, J.S. & Cotanche, D.A. (1994) *J. Comp. Neurol.* **341**, 50–67.
25. Cotanche, D.A., Messana, E.P. & Ofsie, M.S. (1995) *Hear. Res.* **91**, 148–159.
26. Tsue, T.T., Watling, D.L., Weisleder, P., Coltrera, M.D. & Rubel, E. W (1994) *J. Neurosci.* **14**, 140–152.
27. Warchol, M.E. & Corwin, J.T. (1996) *J. Neurosci.* **16**, 5466–5477.
28. Stone, J.S. & Rubel, E. W (2000) *J. Comp. Neurol.* **417**, 1–16.
29. Bhave, S.A., Stone, J.S., Rubel, E. W & Coltrera, M.D. (1995) *J. Neurosci.* **15**, 4618–4628.
30. Stone, J.S., Choi, Y.S., Woolley, S.M.N., Yamashita, H. & Rubel, E. W (1999) *J. Neurocytol.* **28**, 863–876.
31. Hashino, E., TinHan, E.K. & Salvi, R.J. (1995) *Hear. Res.* **88**, 156–168.
32. Presson, J. C, Smith, T. & Mentz, L. (1995) *J. Neurobiol.* **26**, 579–584.
33. Smeets, M., de Witte, T., van der Lely, N., Raymakers, R. & Muus, P. (1999) *Adv. Exp. Med. Biol.* **457**, 557–565.
34. Lippe, W.R., Westbrook, E.W. & Ryals, B.M. (1991) *Hear. Res.* **56**, 203–210.
35. Roberson, D.W., Kreig, C.S. & Rubel, E. W (1996) *Audit. Neurosci.* **2**, 195–205.
36. Wilkins, H.R., Presson, J.C. & Popper, A.N. (1999) *J. Neurobiol.* **39**, 527–535.
37. Murray, R.C. & Calof, A.L. (1999) *Semin. Cell Dev. Biol.* **10**, 421–431.
38. Mumm, J.S., Shou, J. & Calof, A.L. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11167–11172.
39. Gordon, M.K., Mumm, J.S., Davis, R.A., Holcomb, J.D. & Calof, A.L. (1995) *Mol. Cell. Neurosci.* **6**, 363–379.
40. Calof, A.L. & Chikaraishi, D.M. (1989) *Neuron* **3**, 115–127.
41. Mackay-Sim, A. & Kittel, P. (1991) *J. Neurosci.* **11**, 979–984.
42. DeHamer, M.K., Guevara, J.L., Hannon, K., Olwin, B.B. & Calof, A.L. (1994) *Neuron* **13**, 1083–1097.
43. Warchol, M.E. & Corwin, J.T. (1993) *Hear. Res.* **71**, 28–36.
44. Warchol, M.E. (1995) *NeuroReport* **6**, 981–984.
45. Navaratnam, D.S., Su, H.S., Scott, S.P. & Oberholtzer, J.C. (1996) *Nat. Med.* **2**, 1136–1139.
46. Frenz, D.A., Yoo, H. & Liu, W. (1998) *Acta Otolaryngol.* **118**, 651–659.
47. Tsue, T.T., Oesterle, E.C. & Rubel, E. W (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1584–1588.
48. Ofsie, M.S. & Cotanche, D.A. (1996) *J. Comp. Neurol.* **370**, 281–294.
49. Wang, Y. & Raphael, Y. (1997) *Hear. Res.* **97**, 11–18.
50. Hennig, A.K. & Cotanche, D.A. (1998) *J. Neurosci.* **18**, 3282–3296.
51. Livesey, F.J., O'Brien, J.A., Li, M., Smith, A.G., Murphy, L.J. & Hunt, S.P. (1997) *Nature (London)* **390**, 614–618.
52. Matsui, J. L, Oesterle, E.C., Stone, J.S. & Rubel, E. W (2000) *JARO* **1**, 46–63.
53. Corwin, J.T. & Cotanche, D.A. (1989) *J. Comp. Neurol.* **288**, 529–537.
54. Swanson, G.J., Howard, M. & Lewis, J. (1990) *Dev. Biol.* **137**, 243–257.
55. Ernforth, P., Van De Water, T., Loring, J. & Jaenisch, R. (1995) *Neuron* **14**, 1153–1164.
56. Minichiello, L., Piehl, F., Vazquez, E., Schimmang, T., Hokfelt, T., Represa, J. & Klein, R. (1995) *Development (Cambridge, U.K.)* **121**, 4067–4075.
57. Schimmang, T., Alvarez-Bolado, G., Minichiello, L., Vazquez, E., Giraldez, F., Klein, R. & Represa, J. (1997) *Mech. Dev.* **64**, 77–85.
58. Sobkowicz, H.M., Inagaki, M., August, B.K. & Slapnick, S.M. (1999) *J. Neurocytol.* **28**, 17–38.
59. Pickles, J.O. & van Heumen, W.R. (1997) *Dev. Neurosci.* **19**, 476–487.
60. Umemoto, M., Sakagami, M., Fukazawa, K., Ashida, K., Kubo, T., Senda, T. & Yoneda, Y. (1995) *Cell Tissue Res.* **281**, 435–443.
61. Lee, K.H. & Cotanche, D.A. (1996) *Hear. Res.* **94**, 1–13.
62. Oesterle, E.C., Tsue, T.T. & Rubel, E. W (1997) *J. Comp. Neurol.* **380**, 262–274.
63. Warchol, M.E. (1997) *J. Neurobiol.* **33**, 724–734.
64. Fredelius, L. & Rask-Andersen, H. (1990) *Acta Otolaryngol.* **109**, 76–82.
65. Jones, J.E. & Corwin, J.T. (1993) *J. Neurosci.* **13**, 1022–1034.
66. Roberson, D.W. & Rubel, E. W (1994) *Am. J. Otolaryngol.* **15**, 28–34.
67. Jones, J.E. & Corwin, J.T. (1996) *J. Neurosci.* **16**, 649–662.
68. Nathan, C.F. (1987) *J. Clin. Invest.* **79**, 319–326.
69. Brauchle, M., Angermeyer, K., Hubner, G. & Werner, S. (1994) *Oncogene* **9**, 3199–3204.
70. Warchol, M.E. (1999) *J. Neurocytol.* **28**, 889–900.
71. Cotanche, D.A. & Dopyera, C.E. J. (1990) *Hear. Res.* **46**, 29–40.
72. Gleich, O., Dooling, R.J. & Manley, G.A. (1994) *Hear. Res.* **79**, 123–136.
73. Gleich, O., Dooling, R.J. & Presson, J.C. (1997) *J. Comp. Neurol.* **377**, 5–14.
74. Kil, J., Warchol, M.E. & Corwin, J.T. (1997) *Hear. Res.* **114**, 117–126.
75. Oesterle, C., Strohschein, R., Kohler, M., Schnell, M. & Hennig, J. (2000) *J. Magn. Reson. Imaging* **11**, 312–323.
76. Bray, S. (1998) *Cell* **93**, 499–503.
77. Lewis, J. (1996) *Curr. Opin. Neurobiol.* **6**, 3–10.
78. Adam, J., Myat, A., Le Roux, I., Eddison, M., Henrique, D., Ish-Horowicz, D. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 4645–4654.
79. Stone, J.S. & Rubel, E. W (1999) *Development (Cambridge, U.K.)* **126**, 961–973.
80. Haddon, C., Jiang, Y.J., Smithers, L. & Lewis, J. (1998) *Development (Cambridge, U.K.)* **125**, 4637–4644.
81. Riley, B.B., Chiang, M., Farmer, L. & Heck, R. (1999) *Development (Cambridge, U.K.)* **126**, 5669–5678.
82. Lanford, P.J., Lan, Y., Jiang, R., Lindsell, C., Weinmaster, G., Gridley, T. & Kelley, M.W. (1999) *Nat. Genet.* **21**, 289–292.
83. Gridley, T. (1997) *Mol. Cell. Neurosci.* **9**, 103–108.
84. Go, M.J., Eastman, D.S. & Artavanis-Tsakonas, S. (1998) *Development (Cambridge, U.K.)* **125**, 2031–2040.
85. Johnston, L.A. & Edgar, B.A. (1998) *Nature (London)* **394**, 82–84.
86. Stone, J.S. & Cotanche, D.A. (1992) *J. Cell Sci.* **102**, 671–680.
87. Duckert, L.G. & Rubel, E. W (1993) *J. Comp. Neurol.* **331**, 75–96.
88. Marean, G.C., Burt, J.M., Beecher, M.D. & Rubel, E. W (1993) *Hear. Res.* **71**, 125–136.
89. Duckert, L.G. & Rubel, E. W (1990) *Hear. Res.* **48**, 161–182.
90. Ryals, B.M., Westbrook, E.W., Stoots, S. & Spencer, R.F. (1992) *Exp. Neurol.* **115**, 18–22.
91. Wang, Y. & Raphael, Y. (1996) *Hear. Res.* **97**, 11–18.
92. Ofsie, M.S., Hennig, A.K., Messana, E.P. & Cotanche, D.A. (1997) *Hear. Res.* **113**, 207–223.
93. Whitehead, M.C. & Morest, D.K. (1985) *Neuroscience* **14**, 255–276.
94. Whitehead, M.C. & Morest, D.K. (1985) *Neuroscience* **14**, 277–300.
95. Molea, D., Stone, J.S. & Rubel, E. W (1999) *J. Comp. Neurol.* **406**, 183–198.
96. Raphael, Y. (1992) *J. Neurocytol.* **21**, 663–671.
97. Sobkowicz, H.M., August, B.K. & Slapnick, S.M. (1996) *Acta Otolaryngol.* **116**, 257–262.
98. Zheng, J.L., Keller, G. & Gao, W.Q. (1999) *J. Neurosci.* **19**, 2161–2170.
99. Zheng, J.L., Helbig, C. & Gao, W. (1997) *J. Neurosci.* **17**, 216–226.
100. Zheng, J.L., Lewis, A.K. & Gao, W.Q. (1998) *Hear. Res.* **117**, 13–23.
101. Lawlor, P., Marcotti, W., Rivolta, M.N., Kros, C.J. & Holley, M.C. (1999) *J. Neurosci.* **19**, 9445–9458.
102. Bartolami, S., Goodyear, R. & Richardson, G. (1991) *J. Comp. Neurol.* **314**, 777–788.
103. Richardson, G.P., Bartolami, S. & Russell, I.J. (1990) *J. Cell Biol.* **110**, 1055–1066.
104. Lee, K.H. & Cotanche, D.A. (1996) *Audiol. Neurootol.* **1**, 41–53.
105. Hasson, T., Gillespie, P.G., Garcia, J.A., MacDonald, R.B., Zhao, Y., Yee, A.G., Mooseker, M.S. & Corey, D.P. (1997) *J. Cell Biol.* **137**, 1287–1307.
106. Oesterle, E.C., Lurie, D.I. & Rubel, E. W (1997) *J. Comp. Neurol.* **379**, 603–616.

107. Goodyear, R. & Richardson, G. (1999) *J. Neurosci.* **19**, 3761–3772.
108. Kruger, R.P., Goodyear, R.J., Legan, P.K., Warchol, M.E., Raphael, Y., Cotanche, D.A. & Richardson, G.P. (1999) *J. Neurosci.* **19**, 4815–4827.
109. Goodyear, R., Killick, R., Legan, P.K. & Richardson, G.P. (1996) *Hear. Res.* **96**, 167–178.
110. Coutinho, P., Goodyear, R., Legan, P.K. & Richardson, G.P. (1999) *Hear. Res.* **130**, 62–74.
111. Steyger, P.S., Burton, M., Hawkins, J.R., Schuff, N.R. & Baird, R.A. (1997) *Int. J. Dev. Neurosci.* **15**, 417–432.
112. Fekete, D.M., Muthukumar, S. & Karagozeos, D. (1998) *J. Neurosci.* **18**, 7811–7821.
113. Jan, Y.N. & Jan, L.Y. (1998) *Nature (London)* **392**, 775–778.
114. Fekete, D.M. (1996) *Curr. Opin. Neurobiol.* **6**, 533–541.
115. Whitfield, T., Haddon, C. & Lewis, J. (1997) *Semin. Cell Dev. Biol.* **8**, 239–247.
116. Vogel, G. (1998) *Science* **282**, 1235.
117. Lindsell, C.E., Boulter, J., diSibio, G., Gossler, A. & Weinmaster, G. (1996) *Mol. Cell. Neurosci.* **8**, 14–27.
118. Myat, A., Henrique, D., Ish-Horowitz, D. & Lewis, J. (1996) *Dev. Biol.* **174**, 233–247.
119. Dornseifer, P., Takke, C. & Campos-Ortega, J.A. (1997) *Mech. Dev.* **63**, 159–171.
120. Appel, B. & Eisen, J.S. (1998) *Development (Cambridge, U.K.)* **125**, 371–380.
121. Hamada, Y., Kadokawa, Y., Okabe, M., Ikawa, M., Coleman, J.R. & Tsujimoto, Y. (1999) *Development (Cambridge, U.K.)* **126**, 3415–3424.
122. Lewis, A.K., Frantz, G.D., Carpenter, D.A., de Sauvage, F.J. & Gao, W.Q. (1998) *Mech. Dev.* **78**, 159–163.
123. Morrison, A., Hodgetts, C., Gossler, A., Hrabe de Angelis, M. & Lewis, J. (1999) *Mech. Dev.* **84**, 169–172.
124. Robey, E. (1997) *Curr. Opin. Genet. Dev.* **7**, 551–557.
125. Maden, M. (1998) *Subcell. Biochem.* **30**, 81–111.
126. Kelley, M.W., Xu, X.M., Wagner, M.A., Warchol, M.E. & Corwin, J.T. (1993) *Development (Cambridge, U.K.)* **119**, 1041–1053.
127. Ylikoski, J., Pirvola, U. & Eriksson, U. (1994) *J. Comp. Neurol.* **349**, 596–602.
128. Romand, R., Sapin, V. & Dolle, P. (1998) *J. Comp. Neurol.* **393**, 298–308.
129. Raz, Y. & Kelley, M.W. (1999) *Dev. Biol.* **213**, 180–193.
130. LeFebvre, P.P., Malgrange, B., Staecker, H., Moonen, G. & Van De Water, T.R. (1993) *Science* **260**, 692–695.
131. Chardin, S. & Romand, R. (1995) *Science* **267**, 707–711.
132. Chardin, S. & Romand, R. (1997) *Int. J. Dev. Neurosci.* **15**, 497–507.
133. Watt, F.M. & Hogan, B.L. (2000) *Science* **287**, 1427–1430.
134. Corey, D.P. & Breakefield, X.O. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 433–436.
135. Rivolta, M.N. (1997) *Audiol Neurootol* **2**, 36–49.
136. Bermingham, N.A., Hassan, B.A., Price, S.D., Vollrath, M.A., Ben-Arie, N., Eatock, R.A., Bellen, H.J., Lysakowski, A. & Zoghbi, H.Y. (1999) *Science* **284**, 1837–1841.
137. Zheng, J.L. & Gao, W.Q. (2000) *Nat. Neurosci.* **3**, 580–586.
138. Tomarev, S. I., Sundin, O., Banerjee-Basu, S., Duncan, M.K., Yang, J.M. & Piatigorsky, J. (1996) *Dev. Dyn.* **206**, 354–367.
139. Gao, F.B., Brenman, J.E., Jan, L.Y. & Jan, Y.N. (2000) *Genes Dev.* **13**, 2549–2561.
140. Manning, L. & Doe, C.Q. (1999) *Development (Cambridge, U.K.)* **126**, 2063–2071.
141. Rhyu, M.S., Jan, L.Y. & Jan, Y.N. (1994) *Cell* **76**, 477–491.
142. Uemura, T., Shepherd, S., Ackerman, L., Jan, L.Y. & Jan, Y.N. (1989) *Cell* **58**, 349–360.
143. Zhong, W., Feder, J.N., Jiang, M.M., Jan, L.Y. & Jan, Y.N. (1996) *Neuron* **17**, 43–53.
144. Verdi, J.M., Schmandt, R., Bashirullah, A., Jacob, S., Salvino, R., Craig, C.G., Program, A.E., Lipshitz, H.D. & McGlade, C.J. (1996) *Curr. Biol.* **6**, 1134–1145.
145. Verdi, J.M., Bashirullah, A., Goldhawk, D.E., Kubu, C.J., Jamali, M., Meakin, S.O. & Lipshitz, H.D. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 10472–10476.
146. Wakamatsu, Y., Maynard, T.M., Jones, S.U. & Weston, J.A. (1999) *Neuron* **23**, 71–81.
147. Frise, E., Knoblich, J.A., Younger-Shepherd, S., Jan, L.Y. & Jan, Y.N. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11925–11932.
148. Guo, M., Jan, L.Y. & Jan, Y.N. (1996) *Neuron* **17**, 27–41.
149. Spana, E.P. & Doe, C.Q. (1996) *Neuron* **17**, 21–26.
150. Goodyear, R. & Richardson, G. (1997) *J. Neurosci.* **17**, 6289–6301.
151. Lang, H., Bever, M.M. & Fekete, D.M. (2000) *J. Comp. Neurol.* **417**, 205–220.
152. Baird, R.A., Steyger, P.S. & Schuff, N.R. (1996) *Ann. N.Y. Acad. Sci.* **781**, 59–70.
153. Adler, H.J. & Raphael, Y. (1996) *Neurosci. Lett.* **205**, 17–20.
154. Ruben, R.J. (1969) *Laryngoscope* **79**, 546–556.
155. Katayama, A. & Corwin, J.T. (1989) *J. Comp. Neurol.* **281**, 129–135.
156. Lowenheim, H., Furness, D.N., Kil, J., Zinn, C., Gultig, K., Fero, M.L., Frost, D., Gummer, A.W., Roberts, J.M., Rubel, E.W., et al. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4084–4088.
157. Chen, P. & Segil, N. (1999) *Development (Cambridge, U.K.)* **126**, 1581–1590.
158. Lambert, P.R. (1994) *Laryngoscope* **104**, 701–718.
159. Yamashita, H. & Oesterle, E.C. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3152–3155.
160. Kuntz, A.L. & Oesterle, E.C. (1998) *Otolaryngol.-Head Neck Surgery* **118**, 816–824.
161. Baird, R. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11722–11729.

HAIR CELL RECOVERY IN MITOTICALLY BLOCKED CULTURES OF THE BULLFROG SACCCULE

Richard A. Baird*, Miriam D. Burton, David S. Fashena, and Rebecca A. Naeger

Fay and Carl Simons Center for Biology of Hearing and Deafness, Central Institute for the Deaf, 4560 Clayton Road, St. Louis, MO 63110

Hair cells in many nonmammalian vertebrates are regenerated by the mitotic division of supporting cell progenitors and the differentiation of the resulting progeny into new hair cells and supporting cells. Recent studies have shown that nonmitotic hair cell recovery after aminoglycoside-induced damage can also occur in the vestibular organs. Using hair cell and supporting cell immunocytochemical markers, we have used confocal and electron microscopy to examine the fate of damaged hair cells and the origin of immature hair cells after gentamicin treatment in mitotically blocked cultures of the bullfrog saccule. Extruding and fragmenting hair cells, which undergo apoptotic cell death, are replaced by scar formations. After losing their bundles, sublethally damaged hair cells remain in the sensory epithelium for prolonged periods, acquiring supporting cell-like morphology and immunoreactivity. These modes of damage appear to be mutually exclusive, implying that sublethally damaged hair cells repair their bundles. Transitional cells, coexpressing hair cell and supporting cell markers, are seen near scar formations created by the expansion of neighboring supporting cells. Most of these cells have morphology and immunoreactivity similar to that of sublethally damaged hair cells. Ultrastructural analysis also reveals that most immature hair cells had autophagic vacuoles, implying that they originated from damaged hair cells rather than supporting cells. Some transitional cells are supporting cells participating in scar formations. Supporting cells also decrease in number during hair cell recovery, supporting the conclusion that some supporting cells undergo phenotypic conversion into hair cells without an intervening mitotic event.

Hair cells (HCs), the sensory receptors of the vertebrate inner ear, are neuroepithelial cells characterized by apical bundles composed of a single true cilium—the kinocilium—and numerous actin-filled stereocilia (1, 2). These mechanosensory cells are lost with age and are highly susceptible to damage from noise, disease, and ototoxic drugs. This age and trauma-induced susceptibility is a profound health problem because a significant proportion of the human population suffers from deafness or balance disorders resulting directly from HC loss.

Until recently, it was thought that mammalian HCs were produced only during embryonic development and, once lost in mature animals, were gone forever (refs. 3 and 4; but see refs. 5 and 6). We now know that the auditory and vestibular organs of many nonmammalian vertebrates, including fish (7–9), amphibians (10–13), and birds (14–17), produce HCs at a low level throughout life. More importantly, these vertebrates retain the capacity to up-regulate their rate of HC production after acoustic (18–22) and aminoglycoside-induced (23–29) damage to existing HCs. This ability is not retained or is present only to a limited extent in mammals (30–34).

Mitotic HC Regeneration

It has been assumed that HCs and supporting cells (SCs) are derived from a common pool of mitotic precursors during mitotic regeneration (35). This assumption, while consistent with developmental studies demonstrating that HCs and SCs derive from a common progenitor (36), has not been tested by direct observation except in amphibian neuromasts (37–39). There is, however, good evidence that HC production in nonmammalian vertebrates is accomplished by the renewed proliferation of nonsensory SCs. SCs in the auditory (21, 40–43) and vestibular (44) organs leave growth arrest and reenter the G₁-phase of the cell cycle after acoustic or aminoglycoside-induced damage to existing HCs. It remains unclear whether all SCs are capable of serving as HC progenitors. SCs in damaged and undamaged regions reenter the G₁-phase of the cell cycle (45, 46), but only 15% of SCs progress from G₁-phase to S-phase after HC damage, and these cells are limited to the immediate vicinity of the damaged region (47). Proliferating SCs in the fish saccule have no immunocytochemical (48) or ultrastructural (49) features that distinguish them from nonproliferating SCs.

The cellular events that stimulate HC progenitors to divide are only partially understood. Progenitor proliferation is upregulated by diffusible factors released by damaged inner ear organs (50, 51). Progenitors also continue to divide in serumfree medium, indicating that mitotic substances are intrinsic to the inner ear organs (52, 53). A number of fibroblast growth factors that bind and activate tyrosine kinase receptors are known to be present in inner ear organs (54, 55). The expression levels of many of these factors are changed after HC damage, although their mitogenic properties have not yet been determined. Leukocytes, which release growth factors and cytokines, are also likely to play a role in progenitor proliferation (38, 39, 56). The signal transduction pathways downstream of activated tyrosine kinase receptors are less well understood. Recent studies indicate, however, that activation of cyclic adenosine monophosphate (cAMP) leads to increased progenitor proliferation[†] (57) and that many signaling molecules in the cAMP pathway are expressed in the inner ear organs.[‡]

The daughter cells resulting from progenitor division have been presumed to differentiate as new HCs and SCs (19, 20). However, the sequences of events that follow progenitor division, including proliferation, down-regulation, and fate determination, are only beginning to be revealed (refs. 59 and 60; for recent reviews, see refs. 6 and 22). Although it is usually assumed that these postmitotic events mirror those that occur during normal development, this assumption has not been rigorously tested. Many disparities in patterning and organization exist between developing and regenerating inner ear organs, suggest-

*To whom reprint requests should be addressed. E-mail: rbaird@cid.wustl.edu.

[†]Bell, T., Bot, G. & Oberholtzer, J.C. (1999) *Assoc. Res. Otolaryngol. Abstr.* 22, 130 (abstr.).

[‡]Bell, T., Matuai, H., Caruso, L. & Oberholtzer, J.C. (2000) *Assoc. Res. Otolaryngol. Abstr.* 23, 240 (abstr.).

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: C, control; GT, gentamicin-treated; MBC, mitotically blocked C; MBGT, mitotically blocked GT; HC, hair cell; SC, supporting cell; TC, transitional cell.

ing that the genes that regulate these events are not expressed or that their expression is not properly coordinated.

The fate of regenerated cells depends on the activity of transcription factors and other cell fate determinants. Although many transcription factors are expressed in the developing otocyst [see Brigande et al. (61) and Cantos et al. (62), this issue], there is limited evidence that any of them are critical for fate determination in the inner ear organs. Two notable exceptions are the helix-loop-helix transcription factor Math1 and the POU4 transcription factor Brn-3c, both of which are critical for HC development (63, 64). In knockout mice, Math1 deletion prevents HC production in the inner ear organs (63). HCs are transiently produced in Brn-3c knockout mice but do not develop bundles or migrate to their proper epithelial positions, suggesting that this gene plays an important role in HC differentiation, maintenance, or survival (64). The role of these and other genes in postembryonic HC production remains to be determined.

Contact-mediated signaling may play an important role in controlling cell fate during development and regeneration. Notch, an extracellular receptor activated by ligands on adjacent cells, is known to decrease proliferation and regulate pattern formation among sensory cells through lateral inhibition. This receptor, which has been extensively studied in the developing *Drosophila* nervous system, plays an important role in boundary formation and cell lineage decisions (reviewed in refs. 65 and 66). Studies of the developing (67–69) and regenerating (70) inner ear also implicate a role for Notch signaling in HC production. The role that Notch signaling may play in HC development and mitotic HC regeneration is discussed more fully in this issue in papers by Eddison et al. (71) and Stone and Rubel (72).

Nonmitotic HC Regeneration

Recent studies in regenerating nonmammalian (47, 73) and mammalian (32–34) inner ear organs have noted that progenitor proliferation, as determined by BrdUrd immunocytochemistry or tritiated thymidine autoradiography, is limited and that immature HCs in these organs are seldom labeled with immunocytochemical mitotic markers. These discrepancies remain true even when these markers are made continuously available to potential progenitor cells in inner ear organs (47). HC recovery is also sustained after acoustic (74) or aminoglycoside-induced (75) damage in inner ear organs in the presence of mitotic inhibitors such as arabinoside-C and in organotypic cultures under mitotically blocked conditions (73, 76). Taken together, these experiments indicate that HC regeneration can also occur by nonmitotic cellular mechanisms.

The cellular mechanism(s) underlying nonmitotic HC recovery are not as well understood as those underlying mitotic HC recovery. Sobkowicz *et al.* (77) were the first to suggest that damaged HCs could repair their bundles, demonstrating that cochlear HCs could withdraw from the luminal surface and survive the loss of their apical structures. In later experiments, they demonstrated that laser-damaged HCs could regrow their bundles on regaining contact with the luminal surface (78). It has recently been suggested that gentamicin-damaged HCs in the mammalian inner ear may be capable of a similar form of self-repair (34).

The possibility that some SCs might have the capacity to transdifferentiate, i.e., convert phenotypically into HCs without an intervening mitotic event, was first suggested by Jones and Corwin (39). In these experiments, Jones and Corwin observed the development of HC characteristics in SCs that had not undergone a nuclear division after laser-ablating HCs in the axolotl lateral line. Additional morphological evidence for SC conversion in nonmammalian (73–76) and mammalian (79, 80) inner ear organs has come from a number of other *in vivo* studies. In organotypic cultures of the bullfrog saccule, single SCs adjacent to damaged HCs up-regulate calcium-binding proteins after gentamicin ototoxicity, consistent with this hypothesis (76). There is also an *in vivo* loss of SCs in the bullfrog saccule (73) during HC recovery, suggesting that some SCs are converted to HCs during this process.

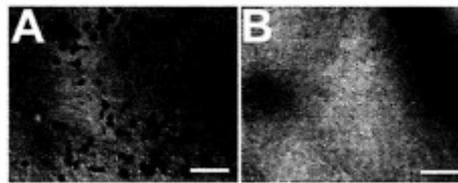


Fig. 1. Micrographs of GT (A) and MBGT (B) cultures incubated for 7 days in modified Wolfe-Quimby culture medium supplemented with 1 μ M BrdUrd. Cultures were fixed for 2 h in 4% paraformaldehyde, permeabilized for 30 min in 1 M HCl containing 0.2% Triton X-100, and blocked for 30 min in PBS containing 10% normal horse serum, 2.5% BSA, and 0.2% Triton X-100. They were then incubated overnight at 4°C with anti-BrdUrd antisera (Caltag, South San Francisco, CA; IU-4) (1:2000), 1 h in biotinylated horse anti-mouse IgG (1:500), and 1 h in avidin-biotin peroxidase complex (1:500). BrdUrd-labeled nuclei were detected with standard diaminobenzidine (DAB) histochemistry. Note the presence of BrdUrd-labeled nuclei in GT (A) and absence of such nuclei in MBGT cultures (B). Bars=25 μ m.

Normal and Mitotically Blocked Cultures of the Bullfrog Saccule

The bullfrog saccule, a sensor of gravity and substrate-borne vibration (10, 81), is a model system for studies of HC transduction (for recent reviews, see refs. 1 and 2). Saccular HCs also increase in number throughout adult life and regenerate after damage, making this organ an ideal system for studying HC development (10, 12, 82) and regeneration (28, 73, 76).

To isolate and study the mechanisms underlying nonmitotic HC recovery, we prepared normal and mitotically blocked saccular cultures by incubating control (C) and gentamicin-treated (GT) saccules in normal medium and medium supplemented with aphidicolin, an inhibitor of DNA polymerase alpha (83). One group of cultures was exposed continuously for 7 days to 1 μ M BrdUrd, a thymidine analogue incorporated into proliferating cells. They were then processed for BrdUrd immunocytochemistry to identify proliferating cells and ascertain the effect of aphidicolin on cell proliferation.

Cell Proliferation in Saccular Cultures

BrdUrd-labeled nuclei in C (data not shown) and GT (Fig. 1A) cultures were found in the sensory macula and the macular margins. Within the sensory macula, the number of BrdUrd-labeled nuclei varied markedly, averaging 500 and 415 in C and GT cultures (Table 1). A small (<3%) number of solitary BrdUrd-labeled nuclei were observed. These nuclei were seen at

Table 1. BrdUrd-labeled nuclei in 7-day MBC and MBGT saccular cultures

Incubation	<i>n</i>	Total nuclei	Solitary nuclei	Paired nuclei
C	4	500.0 \pm 383.3	4.8 \pm 2.2	495.2 \pm 381.4
MBC	4	5.0 \pm 0.8	5.0 \pm 0.8	0.0 \pm 0.0
GT	4	415.0 \pm 477.2	10.7 \pm 5.5	404.3 \pm 471.7
MBGT	4	10.8 \pm 5.6	10.8 \pm 5.6	0.0 \pm 0.0

n, number of cultures; values, means \pm SDs. C and MBC, control cultures incubated in normal (C) and aphidicolin-supplemented (MBC) medium; GT and MBGT, gentamicin-treated cultures incubated in normal (GT) and aphidicolin-supplemented (MBGT) medium.

all epithelial positions and, when located in the apical third of the sensory epithelium, often had mitotic figures, indicating that they were undergoing mitotic division. Most (>97%) BrdUrd-labeled nuclei were found in cell pairs and did not have mitotic figures, implying that they were the progeny of progenitors that had undergone mitotic division.

Aphidicolin markedly decreased cell proliferation in mitotically blocked control (MBC) (data not shown) and gentamicin-treated (MBGT) (Fig. 1B) cultures, reducing the number of BrdUrd-labeled nuclei to <3% of their normal values. Paired BrdUrd-labeled nuclei were eliminated in MBC and MBGT cultures (Table 1), consistent with a blockage of mitotic division. Solitary BrdUrd-labeled nuclei in MBC and MBGT cultures were restricted to the apical third of the sensory epithelium and did not exhibit mitotic figures, implying that they were arrested in the G₂-phase of the cell cycle. Time-lapse studies of Hoechst-labeled MBGT cultures have confirmed that mitotic division did not result from G₂ to M-phase transitions, which would not be detected with BrdUrd immunocytochemistry (R.A.B., unpublished observations).

Morphology and Immunoreactivity in MBC Cultures

A separate group of 11 MBC and 17 MBGT cultures were incubated without BrdUrd up to 9 days to examine the mechanisms of nonmitotic HC recovery. After fixation, these cultures were labeled with phalloidin, a marker of filamentous actin, and immunolabeled with antisera against myosin VI, a known HC marker (84), and a single-chain antibody fragment against a novel cytokeratin expressed in bullfrog SCs (85). The morphology, ultrastructure, and immunolabeling of cells in these cultures were then examined with confocal and transmission electron microscopy.

MBC cultures, as revealed by myosin (red) and cytokeratin (green) immunolabeling, were composed of an organized mosaic of HCs and SCs (Fig. 2A and B). HCs had round apical surfaces with myosin-labeled cell bodies and bundles emerging from a notched cuticular plate. Cytokeratin-labeled SCs, which lacked bundles and cuticular plates, had polygonal apical surfaces. The central region of MBC cultures was a pseudostratified columnar epithelium of cylindrical HCs restricted to the upper two-thirds of the sensory epithelium interspersed with elongated SCs spanning the sensory epithelium. The nuclei of these cells were organized into two layers, with the nuclei of SCs located at or near the basement membrane.

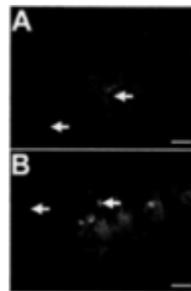


Fig. 2. Luminal surface (A) and mid-epithelium (B) of myosin (red), cytokeratin (green), and Hoechst (blue) labeling a 9-day MBC culture, illustrating sequentially acquired with a confocal laser scanning microscope (Bio-Rad, Radiance 2000). In the central saccule (Bottom, A and B), HCs and SCs are immunolabeled, respectively, with myosin (red) and cytokeratin (green) immunoreactivity. Immature HCs on the macular margin are immunolabeled with both markers (right arrows, A and B), with cytokeratin immunoreactivity in the cytoplasm and perinuclear bodies (left arrows, A and B). Bars=25 μ m.

Mature HCs in the central saccule (Bottom, Fig. 2A and B) had large apical surfaces and short kinocilia with prominent kinociliary bulbs. Immature HCs in the macular margin (Top, Fig. 2A and B) had smaller apical surfaces, long, unbulbed kinocilia, and more elongated cell bodies than mature HCs. Unlike their more mature counterparts, immature HCs had both myosin and cytokeratin immunoreactivity throughout their cell bodies (right arrows) and in prominent perinuclear bodies (left arrows).

Morphology and Immunoreactivity in MBGT Cultures

Gentamicin treatment resulted in widespread damage to saccular HCs, the creation of scar formations, and the appearance of immature HCs in the central saccule. HCs were damaged lethally and sublethally by gentamicin treatment. Some lethally damaged HCs in MBGT cultures were extruded as a unit from the sensory epithelium. This process, although observed with confocal microscopy, was best appreciated by direct observation in MBGT cultures incubated without otolith membranes (R.A.B., unpublished observations). We also observed round epithelial holes in 3- and, to a lesser extent, 5-day MBGT cultures, presumably left by individual HCs extruding from the sensory epithelium.

HCs in MBGT cultures were also fragmented, i.e., broken into an apical remnant, consisting of the bundle and cuticular plate, and a basal remnant, containing the cell nucleus and remaining cytoplasm. HC fragmentation, although seen in the confocal microscope, was best appreciated with transmission electron microscopy (Fig. 3 A and B). Apical HC remnants, which always had intact bundles, were extruded from the sensory epithelium (top arrows). The basal remnants of fragmenting HCs (bottom

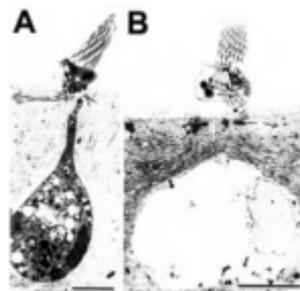


Fig. 3. Transmission electron micrographs (A and B) of apical (top arrows) and basal (filled arrows) remnants of fragmenting HCs in a 7-day MBGT culture. In B, the basal remnant is trapped inside a scar formation formed by the expansion of the apical processes of neighboring SCs. Bars=10 μ m.

arrows) remained in the sensory epithelium, often under the scar formations formed by the expanded processes of neighboring SCs (Fig. 3B). Cytoplasmic remnants of these cells were highly vacuolated, and their nuclei were condensed or disrupted (Fig. 3A), suggesting that they were undergoing apoptotic cell death. Numerous apoptotic bodies were also seen in the sensory epithelium, consistent with this conclusion (A.Lysakowski and R.A.B., unpublished observations). Similar observations have recently been made in cultured explants of the mammalian vestibular organs (89).

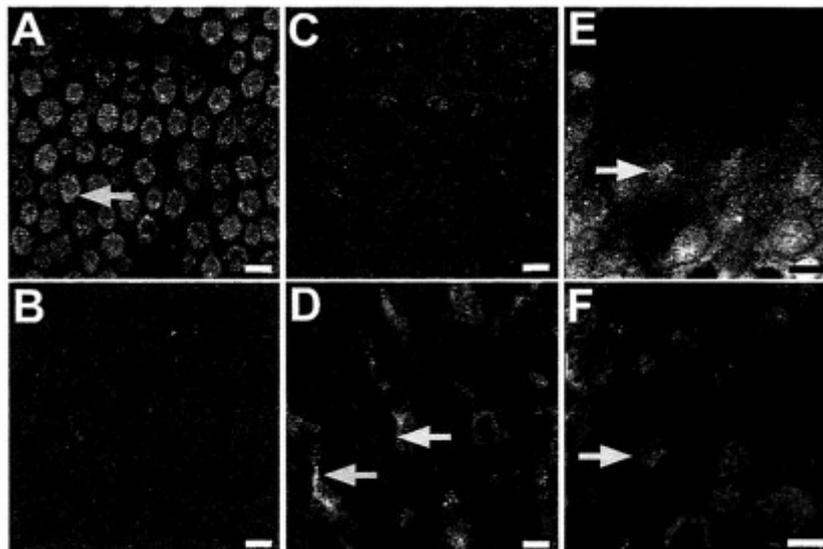


Fig. 4. Luminal surface (A, C, and E) and mid-epithelium (B, D, and F) of 3-day (A and B) and 5-day (C-F) MBGT cultures, illustrating immunolabeling of sublethally damaged HCs. With the exception of their missing bundles, sublethally damaged HCs in 3-day cultures had normal morphology. They also retained their myosin immunoreactivity and, unlike undamaged HCs, had cytokeratin immunoreactivity around their basal bodies and cuticular plates (arrow, A). In 5-day cultures, these cells had degenerating cuticular plates (C), as well as elongated cell bodies, basal nuclei, and basal pseudopodia ending at or near the basement membrane (D). They also had cytokeratin immunoreactivity in their cytoplasm (left arrow, D) and perinuclear bodies (right arrow, D). Sublethally damaged HCs also had enlarged mitochondria and numerous autophagic vacuoles (arrows, E and F). Bars=10 μ m.

The majority of HCs in 3- and 5-day MBGT cultures had partial or missing bundles. These sublethally damaged cells, except for the loss of their bundles, had normal morphology and myosin immunoreactivity and could be differentiated from SCs by their round apical surfaces (Fig. 4A), cuticular plates, and apical nuclei (Fig. 4B). Unlike undamaged HCs, sublethally damaged HCs also had cytokeratin immunoreactivity around their apical and lateral surfaces, particularly near their basal bodies (arrow, Fig. 4A). In 5-day MBGT cultures, sublethally damaged HCs had degenerating cuticular plates and smaller apical surfaces, implying that they were withdrawing from the luminal surface (Fig. 4 C and E). They also had elongated cell bodies, nuclei between the HC and SC nuclear layers, perinuclear bodies (right arrow) and basal pseudopodia (left arrow) ending at or near the basement membrane (Fig. 4D). They also displayed, like immature HCs on the macular margins, both myosin and cytokeratin immunoreactivity (Fig. 4D). Sublethally damaged HCs, whether seen in confocal (Fig. 4F) or transmission electron microscopy (A.Lysakowski and R.A.B., unpublished observations), also had enlarged mitochondria and numerous autophagic vacuoles.

HC Loss and Recovery in MBGT Cultures

To document HC loss and recovery in MBGT cultures, we quantitatively analyzed HC and bundle density in the central saccule at varying time points after gentamicin treatment. Gentamicin treatment induced a rapid loss of bundles, with the density of undamaged bundles in MBGT cultures (filled circles) dropping by 3 days to <5% of the density in MBC cultures (open circles) before returning to \approx 50% of normal levels by 9 days after gentamicin treatment (Fig. 5 Top). HC density in MBGT cultures, determined from the density of myosin-labeled cells, dropped slightly in 3- and 5-day MBGT cultures before returning to its value in MBC cultures (data not shown).

Bundles in MBGT cultures were subdivided into mature (blue bars), damaged (red bars), and immature (yellow bars) bundles (Fig. 5 Middle). Undamaged mature bundle density was low in 3-day MBGT cultures but increased significantly by 9 days after gentamicin treatment, implying that large numbers of hair cells were being repaired. Damaged bundles were seen in large numbers in 3-day MBGT cultures but decreased rapidly at later survival times. The density of immature bundles was small at all survival times, presumably because they rapidly developed into mature bundles.

Transitional Cells and Immature Hair Cells in MBGT Cultures

Scar formations, consisting of four to eight SCs meeting at or near a common vertex, were first seen in 3-day MBGT cultures. These formations, created by the expansion of the apical projections of neighboring SCs into the epithelial spaces vacated by extruding or fragmenting HCs, also had prominent phalloidin-labeled rings, giving them a pie-shaped appearance (Fig. 6A). These rings, composed of actin segments in each of the SCs participating in the scar formation, consisted of a single layer,

suggesting that they were surviving junctional structures from the interface between HCs and their neighboring SCs.

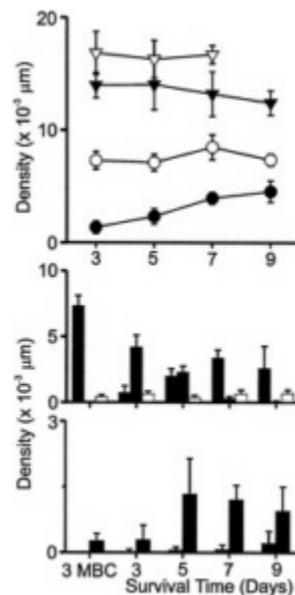


Fig. 5. (Top) Undamaged bundle density (○) and SC density (●) in MBC (open symbols) and MBGT (filled symbols) cultures plotted vs. survival time. (Middle) Density of mature (blue bars), damaged (red bars), and immature (yellow bars) bundles in 3-day MBC cultures (Left) and MBGT cultures (Right) plotted vs. survival time. (Bottom) Density of repairing (blue bars) and nonrepairing (red bars) scars in 3-day MBC cultures (Left) and MBGT cultures (Right) plotted vs. survival time.

Most scar formations showed no signs of repair, i.e., SCs participating in these formations had cytokeratin, but not myosin immunoreactivity and no visible HC characteristics (left arrow, Fig. 6A). In other scar formations, transitional cells (TCs), coexpressing myosin and cytokeratin immunoreactivity, were seen as early as 5 days after gentamicin treatment (right arrow, Fig. 6A). These cells, which we assume to be the precursors of immature HCs, had narrow, elongated cell bodies and nuclei below the HC nuclear layer.

Most (>90%) TCs had round apical surfaces, and they were found between the vertices and outer margins of scar formations (right arrow, Fig. 6A). They also had nuclei between the HC and SC nuclear layers (right middle arrow, Fig. 6B) and basal pseudopodia ending at or near the basement membrane (Fig. 6B and C). Many of these cells also had unusual cell body morphology, traveling horizontally in the basal levels of the sensory epithelium before turning apically to enter scar formations and reach the luminal surface (Fig. 6C). A small (<10%) number of TCs were actually single SCs participating in scar formations. These cells had polygonal or triangular apical surfaces (Fig. 6D), narrow, elongated cell bodies (Fig. 6E and F), and nuclei on or near the basement membrane (lower left arrow, Fig. 6F).

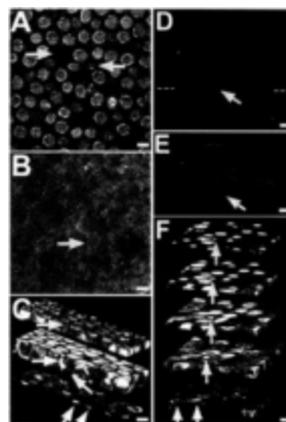


Fig. 6. Luminal surface (A and D), mid-epithelium (B and E), and cut-away views (C and F) of 5-day MBGT cultures, illustrating morphology and immunolabeling of TCs in repairing scar formations. Note round or triangular apical surface (arrows, A and D), elongated cell body (arrows, B and E), and basal nucleus (arrows, C and F). Bars=10 μm.

Immature HCs with long, unbulbed kinocilia and short stereocilia were seen in the central regions of 9-day MBGT cultures (Fig. 7). Unlike TCs, these cells had myosin, but not cytokeratin, immunoreactivity and were not necessarily associated with scar formations. Most of these cells, like TCs, had small round apical surfaces (Fig. 7A and B), nuclei between the HC and SC nuclear layers (middle arrows, Fig. 7C), and basal pseudopodia ending at or near the basement membrane (lower right arrow, Fig. 7C). A few SCs, with triangular apical surfaces and immature bundles, were also seen in scar formations. Immature HCs with more mature bundles, although commonly seen *in vivo*, were harder to document in MBGT cultures.

Scar Formation and SC Loss During HC Recovery

To examine the relationship between scar formation and scar repair, we quantitatively analyzed scar density in the central regions of MBGT cultures at varying time points after gentamicin treatment. Scar density increased 5-fold from its value in MBC cultures by 5 days after gentamicin treatment before leveling off at later survival times (Fig. 5 Bottom). The percentage of repairing scars continued to increase with survival time,

equaling $\approx 15\text{--}20\%$ of all scars in 9-day MBGT cultures. The density of immature HCs was significantly larger than that of repairing scar formations at all survival times, implying that many immature HCs were not associated with repairing scars ($P < 0.01$).

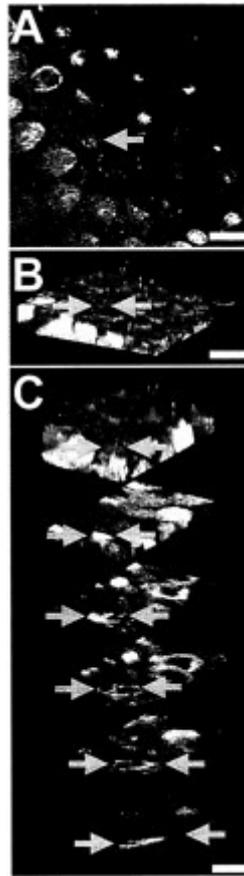


Fig. 7. Luminal surface (A) and cut-away views (B and C) of 9-day MBGT culture, illustrating morphology and immunolabeling of immature HC. Note absence of cytokeratin immunolabeling, round apical surface (arrow, A), kinocilium and immature bundle (arrow, B and C), elongated cell body, basal nucleus (asterisks, B and C), and basal pseudopodia (pointer, C). Bars=10 μm .

Because both undamaged SCs and sublethally damaged HCs expressed cytokeratin immunoreactivity, we also measured SC density to determine whether HC recovery was accompanied by a loss in SC number (Fig. 5 Top). SC density did not change in MBC cultures (open triangles). In MBGT cultures (filled triangles), SC density decreased with increasing survival time, although its value was not significantly different in 3- and 9-day MBGT cultures ($P > 0.05$).

Conclusions

Using immunocytochemical markers at the light and electron microscopic levels, we have shown that it is possible to distinguish damaged HCs from mature HCs and SCs, a task difficult to accomplish solely on morphological grounds because of the cellular disorganization that occurs after gentamicin ototoxicity. Our results indicate that HCs are damaged lethally and sublethally by gentamicin ototoxicity. Lethally damaged HCs are extruded or fragmented; sublethally damaged HCs undergo early bundle loss and withdrawal from the luminal surface, remaining in the sensory epithelium for prolonged periods. We were also able to identify precursor cells expressing HC and SC immunocytochemical markers, a task that has been difficult to accomplish in previous studies because of a lack of cell-specific markers.

Mechanisms of Nonmitotic HC Recovery. At least at the low gentamicin concentration used in this study, our results suggest that HC recovery in MBGT cultures is accomplished by two nonmitotic repair mechanisms, the self-repair of sublethally damaged HCs and the conversion of SCs. We have not examined the responses of MBGT cultures to higher gentamicin concentrations.

Hair Cell Repair. Our observation that large numbers of sublethally damaged HCs survive gentamicin treatment for prolonged periods provides a morphological basis for the self-repair of sublethally damaged HCs. Recent time-lapse studies by us (R.A.B., unpublished observations) confirm that sublethally damaged HCs withdraw from the luminal surface. We have not yet observed by direct observation whether these cells survive to undergo self-repair. Although we have not measured the level of apoptotic cell death in our experiments, we feel that this possibility is unlikely. First, the number of damaged bundles is maximal in 3-day MBGT cultures and decreases markedly at later survival times. A similar result is seen in rat utricles, where gentamicin-induced apoptosis occurs mainly during the two days subsequent to gentamicin treatment (34). Second, extruded HCs were seldom seen by 5 days after gentamicin treatment, and fragmented HCs, unlike sublethally damaged HCs, had intact bundles. These remnants always had intact bundles, implying that HCs with missing bundles did not undergo fragmentation. Taken together, these observations imply that sublethally damaged HCs do not undergo extrusion or fragmentation at later survival times.

Although the origin of most immature HCs could not be determined with certainty in our experiments, three arguments suggest that they derive from sublethally damaged HCs. First, large numbers of sublethally damaged HCs are found in the sensory epithelium before the first appearance of TCs. Second, the morphological, ultrastructural, and immunocytochemical properties of sublethally damaged HCs are similar to those of TCs and immature HCs. Third, TCs and immature HCs have autophagic vacuoles indicative of cellular damage, indicating that they derived from damaged HCs rather than undamaged SCs. We therefore believe that most, if not all, sublethally damaged HCs undergo self-repair and that HC recovery in MBGT cultures is largely accomplished through this mechanism.

Supporting Cell Conversion. Our previous *in vivo* (73) and *in vitro* (73, 76) studies indicated that HC recovery in MBGT cultures was accompanied by morphological changes in single SCs, suggesting that nonmitotic HC regeneration was accomplished by the phenotypic conversion of SCs. These studies also indicated that single SCs participating in scar formations down-regulated cytokeratin, a known SC marker (87), and upregulated calcium-binding proteins (76) known to be expressed in HCs. These immunocytochemical changes appeared to presage the development of morphological HC characteristics, including immature bundles and cuticular plates, in these SCs.

Our current results confirm that a small number of SCs participating in scar formations coexpress HC and SC immunocytochemical markers or display morphological HC characteristics. These morphological and immunocytochemical results indicate that SC conversion plays a role in HC recovery in MBGT cultures. The other possibility, that a cell other than a SC would join an existing scar formation or form a new scar formation, is not consistent with our *in vivo* and *in vitro* observations. More specifically, we have never seen a HC, damaged or undamaged, with a polygonal apical surface, an actin segment, or sharing a common vertex with neighboring SCs.

SC number also decreases in MBGT cultures after gentamicin treatment, consistent with the conclusion that SCs are lost as they undergo conversion during HC recovery. The loss of SCs, however, is significantly lower than the increase in immature HCs, indicating that this process plays only a minor role in HC recovery. The amount of this loss is also less than that seen *in vivo* (73), suggesting that our *in vivo* and *in vitro* gentamicin treatments are not equivalent or that SC conversion is not supported by our culture conditions.

Fate Commitment by Uncommitted Postmitotic Cells. It is possible that a population of uncommitted postmitotic cells could remain latent in the sensory epithelium after undergoing S-phase and change its developmental fate in response to the death of existing HCs. Kelley and his colleagues (88) have shown that embryonic cochlear cultures have a limited ability to repair damage to existing HCs and that cells retain a latent capacity to commit to the HC phenotype for up to 48 h after HCs begin to differentiate. It is not known whether cells in MBGT cultures have a similar latent capacity and, if so, how long this capacity is retained. These uncommitted cells, if they existed, would not be BrdUrd-labeled in our experiments. Nevertheless, we do not believe that shift in developmental fate plays a major role in HC recovery. Because proliferating cells, with the exception of this initial population, would be effectively blocked from entering G₂-phase, we would expect to see an increase in the number of new HCs by this mechanism only at early survival times. By contrast, the number of new HCs in MBGT cultures continued to increase with increasing survival time.

Morphological and Immunocytochemical Properties of HC Precursors. Our results indicate that sublethally damaged HCs in MBGT cultures acquire cytokeratin immunoreactivity on their lateral surfaces after the loss of their hair bundles. In previous studies, this pattern of cytokeratin immunoreactivity was ascribed to supporting cells (59). Whatever their origin, TCs down-regulate cytokeratin immunoreactivity as they differentiate into immature HCs. Cytokeratins are a family of polypeptides that constitute the largest and most complex class of intermediate filaments (58). In epithelial cells, they form a structural network that spans the cell cytoplasm, linking the plasma membrane, nucleus, and other cytoskeletal elements. This structural network is apparently achieved through a recently identified family of intermediate-filament-associated proteins that form cross-bridges between intermediate filaments and other cytoskeletal elements. Although the exact role of cytokeratin in the inner ear is unclear, detailed studies of its cellular distribution in mitotic and nonmitotic HC precursors should prove helpful in delineating the stages of HC differentiation.

Our results also indicate that the morphology and epithelial position of TCs and immature HCs are different from that of mature HCs. More specifically, the former cells have smaller apical surfaces, more elongated cell bodies, and more basal nuclei than the latter cells. Recent studies using calcium-binding proteins (59) and class III β -tubulin (60) as HC immunocytochemical markers also indicate that mitotic HC precursors have nuclei located well below the luminal surface and appear to have physical connection with both the basement membrane and the luminal surface. These morphological features suggest that mitotic HC precursors migrate from the luminal surface to more basal positions before establishing their permanent positions. This migration may be necessary for mitotic and nonmitotic HC precursors to obtain positional cues for proliferation or differentiation from contact with other cells or the extracellular matrix of the basement membrane.

Maturation of mitotic and nonmitotic HC precursors coincides with a shortening of the basal pseudopodia and an apical migration of the cell nucleus, both of which remove these cells from the influence of the basement membrane. Interestingly, many TCs travel horizontally in the basal levels of the sensory epithelium before turning to reach the luminal surface. They are also found on the margins rather than the vertices of scar formations, implying that they were not able to undergo further differentiation without obtaining patterning signals derived from contact with adjacent SCs. These patterning signals, which may be mediated by Notch and its associated ligands, might explain why multiple TCs were seldom seen in scar formations and ensure the faithful reproduction of the normal mosaic of HCs and SCs in repairing cultures.

Interactions Between Mitotic and Nonmitotic HC Recovery. Our results help to explain the early appearance of immature HCs and to reconcile differences between the number of proliferating and regenerating cells in previous studies (31, 32, 47). They also indicate that HC regeneration, at least in the bullfrog saccule, is accomplished by a combination of mitotic and nonmitotic mechanisms. It is not yet clear whether this is a general property of all auditory and vestibular organs. It is possible that using mitotic and nonmitotic mechanisms of HC regeneration allows inner ear organs to respond to ongoing and injury-induced HC loss with a higher degree of flexibility than could be attained by either mechanism alone. Our *in vivo* and *in vitro* results imply that nonmitotic HC regeneration creates new HCs more rapidly than mitotic HC regeneration (73). Nonmitotic mechanisms of HC regeneration might therefore be more suitable for producing larger numbers of HCs after catastrophic HC loss, where rapid HC production might be critical for preventing the subsequent degeneration of afferent neurons. Nonmitotic mechanisms of HC regeneration can also produce new HCs without the metabolic cost of maintaining a stem cell population or the cellular disruption associated with mitotic division. Unfortunately, nonmitotic mechanisms of HC regeneration, unlike mitotic division, are limited in their ability to restore HC and SC density to normal levels. Thus, a combination of mitotic and nonmitotic mechanisms of HC regeneration may normally be required to bring about full recovery.

Time-lapse experiments, using vital markers against the plasma membrane, cytoplasmic organelles, and ion channels and the immunocytochemical markers used in the present study, are underway in our laboratory to confirm the fate of sublethally damaged HCs and the origin of immature HCs in MBGT cultures (R.A.B., unpublished observations). Similar time-lapse experiments are being pursued by other laboratories[§] These studies should provide much needed knowledge about the development of mitotic and nonmitotic HC precursors, including new stage-specific markers correlated with particular morphological and physiological characteristics and should lead to a better understanding of the basic mechanisms underlying proliferation and differentiation in these important sensory cells.

We thank Drs. Mark Warchol and Dwayne Simmons for helpful comments on this manuscript. Funding for this research was provided by National Institutes of Health Grant DC02040 and National Aeronautics and Space Administration Grant NCC 2-651.

1. Hudspeth, A.J. (1989) *Nature (London)* **341**, 397-404.

[§]Gale, J.E. & Corwin, J.T. (1997) *Soc. Neurosci. Abstr.* 23, 1822 (abstr.).

[¶]Meyers, J. & Corwin, J.T. (2000) *Assoc. Res. Otolaryngol. Abstr.* 23, 239 (abstr.).

2. Hudspeth, A.J. (1997) *Neuron* **19**, 947–950.
3. Corwin, J.T. & Warchol, M.E. (1991) *Annu. Rev. Neurosci.* **14**, 301–333.
4. Corwin, J.T., Warchol, M.E. & Kelley, M.T. (1993) *Curr. Opin. Neurobiol.* **3**, 32–37.
5. Staecker, H. & Van De Water, T. (1998) *Curr. Opin. Neurobiol.* **8**, 480–487.
6. Stone, J.S., Oesterle, E.C. & Rubel, E.W. (1998) *Curr. Opin. Neurol.* **11**, 17–24.
7. Corwin, J.T. (1981) *J. Comp. Neurol.* **20**, 541–553.
8. Popper, A.N. & Hoxter, B.H. (1984) *Hear. Res.* **15**, 133–142.
9. Popper, A.N. & Hoxter, B.H. (1990) *Hear. Res.* **45**, 33–40.
10. Lewis, E.R. & Li, C.W. (1975) *Brain Res.* **83**, 35–50.
11. Li, C.W. & Lewis, E.R. (1979) *Ann. Otol. Rhinol. Laryngol.* **88**, 427–437.
12. Corwin, J.T. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3911–3915.
13. Kelley, M.W., Ochiai, C.K. & Corwin, J.T. (1992) *Hear. Res.* **59**, 108–115.
14. Jorgensen, J.M. & Mathiesen, C. (1988) *Naturwissenschaften* **75**, 319–320.
15. Ryals, B.M. & Rubel, E.W. (1988) *Science* **240**, 1774–1776.
16. Katayama, A. & Corwin, J.T. (1989) *J. Comp. Neurol.* **281**, 129–135.
17. Ryals, B.M. & Westbrook, E.W. (1990) *Hear. Res.* **50**, 87–96.
18. Cotanche, D.A. (1987) *Hear. Res.* **30**, 181–196.
19. Corwin, J.T. & Cotanche, D.A. (1988) *Science* **240**, 1772–1774.
20. Ryals, B. & Rubel, E.W. (1988) *Science* **240**, 1774–1776.
21. Girod, D.A., Duckert, L.G. & Rubel, E.W. (1989) *Hear. Res.* **42**, 175–194.
22. Cotanche, D.A., Lee, K.H., Stone, J.S. & Picard, D.A. (1994) *Anat. Embryol.* **189**, 1–18.
23. Cruz, R.M., Lambert, P.R. & Rubel, E.W. (1987) *Arch. Otolaryngol. Head Neck Surg.* **113**, 1058–1062.
24. Girod, D.A., Tucci, D.L. & Rubel, E.W. (1991) *Laryngoscope* **101**, 1139–1149.
25. Hashino, E., Tanaka, Y. & Sokabe, M. (1991) *Hear. Res.* **52**, 356–368.
26. Lippe, W.R., Westbrook, E.W. & Ryals, B.M. (1991) *Hear. Res.* **56**, 203–210.
27. Yan, H.Y., Saidel, W.M., Chang, J.S., Presson, J.C. & Popper, A.N. (1991) *Proc. R. Soc. London Ser. B* **245**, 133–138.
28. Baird, R.A., Torres, M.A. & Schuff, N.R. (1993) *Hear. Res.* **65**, 164–174.
29. Weisleder, P. & Rubel, E.W. (1993) *J. Comp. Neurol.* **331**, 97–110.
30. Forge, A., Li, L., Corwin, J.T. & Nevill, G. (1993) *Science* **259**, 1616–1619.
31. Warchol, M.E., Lambert, P.R., Goldstein, B.J., Forge, A. & Corwin, J.T. (1993) *Science* **259**, 1619–1622.
32. Rubel, E.W., Dew, L.A. & Roberson, D.W. (1995) *Science* **267**, 701–703.
33. Zine, A. & de Ribaupierre, F. (1998) *NeuroReport* **9**, 263–268.
34. Zheng, J.L., Keller, G. & Gao, W.-Q. (1999) *J. Neurosci.* **19**, 2161–2170.
35. Corwin, J.T., Jones, J.E., Katayama, A., Kelley, M.T. & Warchol, M.E. (1991) in *Regeneration of Vertebrate Sensory Receptor Cells*, eds. Bock, G.R. & Whelan, J. (Wiley, Chichester), pp. 103–120.
36. Fekete, D.M., Muthukumar, S. & Karageos, D. (1998) *J. Neurosci.* **18**, 7811–7821.
37. Balak, K.J., Corwin, J.T. & Jones, J.E. (1990) *J. Neurosci.* **10**, 2502–2512.
38. Jones, J.E. & Corwin, J.T. (1993) *J. Neurosci.* **13**, 1022–1034.
39. Jones, J.E. & Corwin, J.T. (1996) *J. Neurosci.* **16**, 649–662.
40. Raphael, Y. (1992) *J. Neurocytol.* **21**, 663–671.
41. Hashino, E. & Salvi, R.J. (1993) *J. Cell Sci.* **105**, 23–31.
42. Stone, J.S. & Cotanche, D. (1994) *J. Comp. Neurol.* **341**, 50–67.
43. Gleich, O., Dooling, R.J. & Presson, J.C. (1997) *J. Comp. Neurol.* **377**, 5–14.
44. Tsue, T.T., Watling, D.L., Weisleder, P., Coltrera, M.D. & Rubel, E.W. (1994) *J. Neurosci.* **14**, 140–152.
45. Bhave, S.A., Stone, J.S., Rubel, E.W. & Coltrera, M. (1995) *J. Neurosci.* **5**, 4618–4628.
46. Presson, J.C., Smith, T. & Mentz, L. (1995) *J. Neurobiol.* **26**, 579–584.
47. Roberson, D.W., Krieg, C.S. & Rubel, E.W. (1996) *Aud. Neurosci.* **2**, 195–205.
48. Presson, J.C. (1994) *Hear. Res.* **80**, 1–9.
49. Presson, J.C., Lanford, P.J. & Popper, A.N. (1996) *Hear. Res.* **100**, 10–20.
50. Tsue, T.T., Oesterle, E.C. & Rubel, E.W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1584–1588.
51. Warchol, M.E. & Corwin, J.T. (1996) *J. Neurosci.* **16**, 5466–5477.
52. Warchol, M.E. & Corwin, J.T. (1993) *Hear. Res.* **71**, 28–36.
53. Warchol, M.E. (1995) *NeuroReport* **6**, 981–984.
54. Lee, K. & Cotanche, D.A. (1996) *Hear. Res.* **94**, 1–13.
55. Pickles, J.O. & van Heumen, W.R. (1997) *Dev. Neurosci.* **19**, 476–487.
56. Warchol, M.E. (1997) *J. Neurobiol.* **33**, 724–734.
57. Navaratnam, D.S., Su, H.S., Scott, S.P. & Oberholtzer, J.C. (1996) *Nat. Med.* **2**, 1136–1139.
58. Houseweart, M.K. & Cleveland, D.W. (1998) *Curr. Opin. Cell Biol.* **10**, 93–101.
59. Stone, J.S., Leano, S.G., Baker, L.P. & Rubel, E.W. (1996) *J. Neurosci.* **16**, 6157–6174.
60. Stone, J.S. & Rubel, E.W. (2000) *J. Comp. Neurol.* **417**, 1–16.
61. Brigande, J.V., Kiernan, A.E., Gao, X., Iten, L.E. & Fekete, D.M. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11700–11706.
62. Cantos, R., Cole, L.K., Acampora, D., Simeone, A. & Wu, D.K. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11707–11713.
63. Bermingham, N.A., Hassan, B.A., Price, S.D., Vollrath, M.A., Ben-Arie, N., Eatock, R.A., Bellen, H.J., Lysakowski, A. & Zoghbi, H.Y. (1999) *Science* **284**, 1837–1841.
64. Xiang, M., Gao, W.-Q., Hasson, T. & Shin, J.J. (1998) *Development* **125**, 3935–3946.
65. Fekete, D.M. (1996) *Curr. Opin. Neurobiol.* **6**, 533–541.
66. Lewis, J.E. (1996) *Curr. Opin. Neurobiol.* **6**, 3–10.
67. Whitfield, T., Haddon, C. & Lewis, J. (1997) *Semin. Cell. Dev. Biol.* **8**, 239–247.
68. Vogel, G. (1998) *Science* **282**, 1235.
69. Lanford, P.J., Lan, Y., Jiang, R., Lindsell, C., Weinmaster, G., Gridley, T. & Kelley, M.W. (1999) *Nat. Genet.* **21**, 289–292.
70. Stone, J.S. & Rubel, E.W. (1999) *Development* **126**, 961–973.
71. Eddison, M., Le Roux, I. & Lewis, J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11692–11699.
72. Stone, J.S. & Rubel, E.W. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11714–11721.
73. Baird, R.A., Steyger, P.S. & Schuff, N.R. (1996) *Ann. N.Y. Acad. Sci.* **781**, 59–70.
74. Adler, H.J. & Raphael, J. (1996) *Neurosci. Lett.* **210**, 73–76.
75. Adler, H.J., Komeda, M. & Raphael, Y. (1997) *Int. J. Dev. Neurosci.* **15**, 375–385.
76. Steyger, P.S., Burton, M.D., Hawkins, J.R., Schuff, N.R. & Baird, R.A. (1997) *Int. J. Dev. Neurosci.* **15**, 417–432.
77. Sobkowicz, H., August, B.K. & Slapnick, S.M. (1996) *Acta Otolaryngol.* **116**, 257–262.
78. Sobkowicz, H., August, B.K. & Slapnick, S.M. (1997) *Int. J. Dev. Neurosci.* **15**, 463–484.
79. Li, L. & Forge, A. (1997) *Int. J. Dev. Neurosci.* **15**, 433–446.
80. Forge, A., Li, L. & Nevill, G. (1998) *J. Comp. Neurol.* **397**, 69–88.
81. Eatock, R.A., Corey, D.P. & Hudspeth, A.J. (1987) *J. Neurosci.* **7**, 2821–2836.
82. Kelley, M.W. & Corwin, J.T. (1992) *Hear. Res.* **59**, 108–115.
83. Spadari, S., Focher, F., Sala, F., Ciarrocchi, G., Koch, G., Falaschi, A. & Pedrali-Noy, G. (1985) *Arzneimittelforschung* **35**, 1108–1116.
84. Hasson, T., Gillespie, P.G., Garcia, J.A., MacDonald, R.B., Zhao, Y., Yee, A.G., Mooseker, M.S. & Corey, D.P. (1997) *J. Cell Biol.* **137**, 1287–1307.
85. Cyr, J.L., Bell, A.M. & Hudspeth, A.J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 4908–4913.
86. Forge, A. & Li, L. (2000) *Hear. Res.* **139**, 97–115.
87. Meiteles, L.Z. & Raphael, Y. (1994) *Ann. Otol. Rhinol. Laryngol.* **103**, 149–155.
88. Kelley, M.W., Talreja, D.R. & Corwin, J.T. (1995) *J. Neurosci.* **15**, 3013–3026.

TWO MECHANISMS FOR TRANSDUCER ADAPTATION IN VERTEBRATE HAIR CELLS

Jeffrey R. Holt and David P. Corey

Department of Neurobiology, Howard Hughes Medical Institute, Harvard Medical School and Massachusetts General Hospital, Wellman 414, Boston, MA 02114

Deflection of the hair bundle atop a sensory hair cell modulates the open probability of mechanosensitive ion channels. In response to sustained deflections, hair cells adapt. Two fundamentally distinct models have been proposed to explain transducer adaptation. Both models support the notion that channel open probability is modulated by calcium that enters via the transduction channels. Both also suggest that the primary effect of adaptation is to shift the deflection-response $[I(X)]$ relationship in the direction of the applied stimulus, thus maintaining hair bundle sensitivity. The models differ in several respects. They operate on different time scales: the faster on the order of a few milliseconds or less and the slower on the order of 10 ms or more. The model proposed to explain fast adaptation suggests that calcium enters and binds at or near the transduction channels to stabilize a closed conformation. The model proposed to explain the slower adaptation suggests that adaptation is mediated by an active, force-generating process that regulates the effective stimulus applied to the transduction channels. Here we discuss the evidence in support of each model and consider the possibility that both may function to varying degrees in hair cells of different species and sensory organs.

The ability to sense sound and head position has proven advantageous for vertebrate animals. For these two purposes evolution has devised a common solution: the mechanosensory hair bundle. This organelle, exquisite in form and function, crowns the apical surface of the sensory hair cells in the inner ears of vertebrates. Hair bundles comprise a staircase array of 20–300 microvilli, called stereocilia, that range in height from a few to tens of micrometers (Fig. 1A). The stereocilia contain a rigid core of cross-linked actin filaments (1–3) and do not bend under normal conditions. Their actin cores taper at the base where they insert into the cuticular plate, thus deflections cause them to pivot instead (4). The stereocilia, ensheathed by the cell membrane, are held together by three sets of extracellular links that extend laterally in a symmetrical manner. A fourth set of links, called tip links, extends from the tip of one stereocilium to the side of its adjacent, taller neighbor along the bundle's axis of symmetry (Fig. 1B and D; ref. 5). The tip links are ≈ 10 nm in diameter and ≈ 150 nm long (6, 7). Deflection of the hair bundle toward the tallest stereocilia gates cation-selective transduction channels (Fig. 1C and D). With a short latency (≈ 10 μ sec at mammalian temperatures) the channels open and generate a receptor current of hundreds of picoamperes that is carried primarily by potassium and calcium ions (8–10). Transduction channels probably are located at either end of the tip links, and based on single-channel conductance estimates of ≈ 100 pS, some hair cells may have >200 functional channels.

Transduction Model

Taken together, the following evidence supports a model for hair cell mechanoelectric transduction that is now widely accepted (Fig. 1D). The short latency of the response argues against activation of a second-messenger cascade and favors direct activation of a mechanically sensitive channel. Bundle deflection parallel to the sensitive axis is posited to stretch elastic "gating

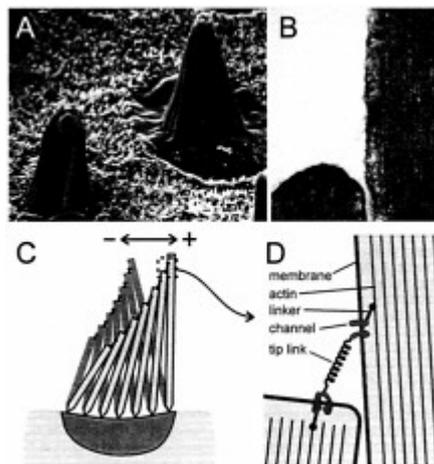


Fig. 1. Ultrastructure and function of sensory hair bundles. (A) Scanning electron micrograph ($\times 2,500$) of two hair bundles in the sensory macula of the bullfrog saccule. The bundles have a morphological axis of polarity defined by the graded heights of the stereocilia (Right) and the eccentric placement of the kinocilium (Left), such that the tallest stereocilia are at the positive edge. These bundles are about $8 \mu\text{m}$ tall and contain 50–60 stereocilia. (B) Transmission electron micrograph ($\times 50,000$) of the tip of a stereocilium and its adjacent, taller neighbor. Note the fine filamentous tip link extending between the two stereocilia. At each end is an electron-dense plaque between the membrane and the actin core. The tip link is 150–200 nm in length; stereocilia are ≈ 400 nm in diameter. (C) Schematic diagram of a hair bundle. Note that the taper at the base of the stereocilia allow them to pivot rather than bend. In this view a bundle deflection to the left is considered a negative stimulus that allows transduction channels to close and a deflection to the right is positive and causes them to open. (D) Schematic diagram illustrating the gating spring model for transduction. The tip link may be the morphological correlate of the gating spring. When the bundle is deflected in the positive direction the tip links stretch, which would in turn pull directly on the channels and cause them to open.

springs," which pull directly on the channels (9). Direct channel gating was supported by experiments of Howard and Hudspeth (11) who found that hair bundle stiffness decreased as the channels opened. This finding implicated a direct gating mechanism in series with the gating springs. The discovery of tip links (5) oriented along the sensitive axis suggested that these fine extracellular filaments are the morphological correlates of the gating spring. Furthermore, Assad *et al.* (6) found that brief exposure to the calcium chelator, 1,2-bis(2-aminophenoxy)ethane-*N,N,N',N'*-tetraacetate (BAPTA), concurrently abolished transduction and tip links. Using the same manipulation, Zhao *et al.* (12) went on to show that if the hair cells were allowed time to recover for 12–24 h, transduction and tip links returned with the same time course. Detailed analysis of quick-freeze, deep-etch electron micrographs of tip-link structure has revealed at least two helical strands with 2–3 branches at their insertion into the stereocilia membrane (13). If there is a transduction channel for each branch, there may be up to six channels per tip link.

Calcium Metabolism

Experiments with Ca^{2+} -sensitive dyes in the patch pipette have demonstrated that transduction channels have a significant Ca^{2+} permeability, channels can be located at either end of tip links, and the concentration rises rapidly in the tips of the stereocilia when channels open (14–16). As discussed in the next section, the Ca^{2+} concentration near the transduction channels has a pronounced feedback effect on the channel open probability. Knowledge of calcium metabolism within the tips of the stereocilia is critical for understanding transduction and adaptation.

Calcium that enters the tips may meet four possible fates: it can diffuse down the stereocilia, bind to mobile buffers that diffuse away, bind to fixed buffers, or be extruded by calcium pumps in the stereocilia membranes. Models indicate that all four mechanisms are required to account for the fluorescence transients that follow transduction channel opening (17). The Ca^{2+} buffers in stereocilia have not yet been identified but may be a protein such as calbindin or calretinin (18, 19), and the effective endogenous buffer concentration is equivalent to 0.1–0.4 mM BAPTA (20). Recent work on Ca^{2+} extrusion from the stereocilia has focused on the plasma membrane Ca^{2+} ATPases (PMCA). Immunocytochemistry and quantitative immunoblotting indicated that the stereocilia may contain up to 2,000 molecules/ μm^2 of a PMCA, enough to account for an outward pump current of several picoamperes (21). Further implicating PMCA is the finding that mice with a mutation or a targeted deletion of the gene for PMCA2 are deaf and have vestibular deficits (22, 23).

Despite the efforts of the hair cell to maintain low levels of Ca^{2+} , the concentration may rise as high as $50 \mu\text{M}$ within close proximity to the transduction channels where it can act to modulate transducer adaptation.

Adaptation

In response to a sustained bundle deflection, the hair cell receptor current declines or adapts, through closure of the transduction channels (Fig. 2A). The decline in current could occur in three ways: by a decrease in slope sensitivity of the channels, so that a larger deflection is needed to open more channels; by an inactivation so that the number of channels available to be opened is reduced; or by shifting the stimulus-response $I(X)$ relationship without a reduction in sensitivity. Hair cell adaptation has been found to be primarily associated with a shift of the $I(X)$ relation, whereby the hair cell moves the relation in the direction of the applied stimulus, thus bringing the channel open probability back toward the resting value (Fig. 2B and C) (9, 24, 25).

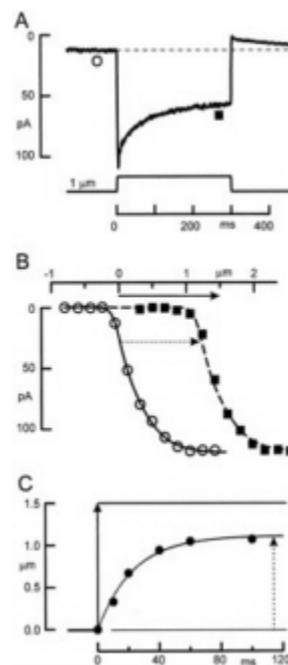


Fig. 2. Properties of adaptation in a mouse utricular type II hair cell. (A) Transduction current evoked by deflecting the hair bundle with a stiff probe mounted on a bimorph stimulator. The current declines in response to a step deflection. When the bundle is returned to its rest position there is an overshoot that eventually returns to its initial level. (B) The $I(X)$ relation shifts in the direction of the applied stimulus. These data were taken from a family of transduction currents (not shown). (C) Magnitude of the $I(X)$ shift as a function of time. The position of $I(X)$ curves such as those in B were measured at several time points after the onset of a $1.5\text{-}\mu\text{m}$ deflection. Fitting these data with an exponential function indicated that the curve shifted with a time constant of 23 msec. Solid arrow indicates magnitude of the stimulus; dashed arrow shows final magnitude of the $I(X)$ shift. Data were replotted from figure 5 in ref. 25.

However, adaptation is not complete: the shift of the $I(X)$ curve is not as large as the sustained bundle deflection that evokes it, so the receptor current does not return all of the way

to the resting level (Fig. 2 A and C). The extent of adaptation is 60–80%, that is, adaptation shifts the I(X) curve by 60–80% of the magnitude of the bundle deflection (25, 26). Thus the major component of the receptor current signals phasic bundle deflections whereas the remaining nonadapted component signals more tonic stimuli.

Calcium inside the tips of the stereocilia has been found to affect both the speed and extent of adaptation considerably. To investigate its role, the concentration in the tips has been manipulated a number of ways, each with measurable effects. A decrease in extracellular Ca^{2+} decreases Ca^{2+} flux through the transduction channels and slows the rate of adaptation, measured either as the rate of shift of the I(X) curve or the decline in receptor current (9, 24, 25, 27–29). An increase in the concentration of a calcium buffer within the cell causes Ca^{2+} to be more rapidly bound once it enters and slows the rate of adaptation (16, 29). Depolarizing the hair cell reduces the driving force for Ca^{2+} influx, reduces Ca^{2+} entry, and reduces or abolishes adaptation (28).

The precise site of action for Ca^{2+} is not clear but it must be within $\approx 1 \mu\text{m}$ of the transduction channels. If Ca^{2+} entry is reduced by depolarization, adaptation is slowed; upon repolarization, Ca^{2+} enters the bundle and adaptation proceeds within 1–2 msec. Thus, Ca^{2+} must find the adaptation control site very near the transduction channel (28). Furthermore, Ca^{2+} buffers with slow binding kinetics, like EGTA, have little effect on adaptation, whereas BAPTA, which has more rapid kinetics, has a significant effect on adaptation rate and extent (29). The interpretation is that BAPTA binds Ca^{2+} before it has time to diffuse to its site of action (20).

Initial descriptions of adaptation characterized it as one mechanism with a single exponential time course. Time constants were reported that ranged between 3 and 100 s of msec (24, 25, 28, 29). Recent work suggests that adaptation may have two different mechanisms, with one as fast as 0.3 msec (30). The latter is more prominent for small deflections, but saturates with larger ones where the slower process dominates. Two distinct models for adaptation have been proposed; one involves a mechanical adjustment of tension on the gating spring (31), and the other involves Ca^{2+} binding directly to a site at or near the channel that alters the relation between tension and open probability (29, 32). In the following sections we consider the evidence for each model and consider the possibility of these acting simultaneously in the hair cell.

The Active Motor Model. This model proposes that in response to a maintained bundle deflection the hair cell actively adjusts the tension in each gating spring to return it toward its resting level (Fig. 3A). Howard and Hudspeth (31) used a flexible glass stimulus probe to measure the mechanical correlates of adaptation in bullfrog saccular hair cells. An applied force caused a rapid bundle deflection that was followed by an additional, slower deflection (Fig. 4). The latter had the same time course as adaptation (≈ 30 msec). This slow deflection occurred for forces parallel to the bundle's axis of sensitivity, but did not occur for forces directed perpendicular to that axis. When adaptation of the receptor current slowed or disappeared during prolonged recording, the slow deflection did as well. Because these bundle relaxations had the same time course as adaptation, Howard and Hudspeth (31) proposed that the mechanical relaxation might be the result of a movement of the upper tip-link attachment point along the side of the stereocilium. Thus, the model suggests that an active motor complex—probably in the vicinity of the upper tip-link insertion—is continuously trying to “climb up” the stereocilium to increase tension in the tip link. In response to a positive deflection, tension in the tip link increases and opens channels. As adaptation proceeds, the motor slips down the side of the stereocilium, tip-link tension decreases, the channels begin to close, and the bundle relaxes forward. In response to a negative deflection, tip-link tension decreases and closes channels. In this case the motor complex climbs up the stereocilium, restores tension to the tip link, and pulls the bundle in the negative direction. At steady state, the rate of motor slipping equals the rate of climbing and the steady-state tension maintains 10–20% of the channels in the open conformation. The slipping and climbing of adaptation motors can be measured as a shift of the I(X) curve in the direction of the applied stimulus. To view this in terms of a shift of the upper attachment point for a single tip link, the geometrical gain of the bundle must be taken into account (≈ 0.12 for bullfrog hair cells). For example, a positive deflection of $0.5 \mu\text{m}$ would stretch the gating spring by about 60 nm, and a subsequent slippage of 48 nm would allow

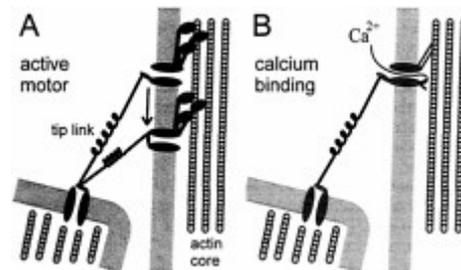


Fig. 3. Schematic diagrams illustrating the two models for hair cell adaptation. (A) The active motor model proposes that when the bundle is deflected in the positive direction, the motor cannot resist the increased tension and slips down the stereocilium, reducing tension and allowing channels to close. In addition, Ca^{2+} entry through open channels accelerates the rate of slipping. Conversely, when the bundle is deflected in the negative direction, the reduced tension allows the motor to climb and restore tension to reopen channels. (B) The calcium-dependent closure mechanism proposes that when the channels open calcium enters the stereocilia and binds to a site on or near the channel protein. Bound Ca^{2+} promotes a closure of the channel. When deflected in the negative direction, the Ca^{2+} concentration falls and reduces the inhibition of the channel.

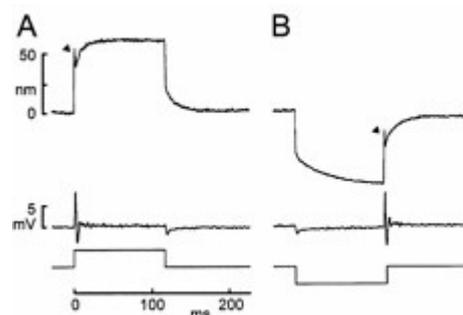


Fig. 4. The “twitch” and subsequent slow relaxations of the hair bundle during maintained force steps. (A) The positive force stimulus in this experiment caused a quick forward deflection of ≈ 50 nm (arrowhead), a rapid twitch back of ≈ 10 nm, and then a relaxation forward of ≈ 25 nm. The receptor potential recorded simultaneously (middle trace) showed a transient depolarization, and then an oscillation during and after the twitch. (B) A negative force evoked a quick negative deflection and slower negative relaxation. The twitch was not observed until the bundle was returned to the rest position (arrowhead) and was associated with an oscillation. [Reprinted with permission from ref. 52. (Copyright 1996, Society for Neuroscience).]

≈80% of the channels to close. Because the cores of the stereocilia are composed of actin of the proper polarity, it has been suggested that a myosin motor could be a suitable mechanoenzyme to power transducer adaptation in hair cells.

This simple model has three basic tenants. First, it suggests that adaptation is a mechanical process that adjusts tension in the physiologically defined gating spring by moving one end. Second, by correlating the gating spring with the tip link, it predicts that the morphologically defined, electron-dense attachment of the tip link also moves during adaptation. Third, it suggests a protein of the myosin family as the leading candidate for the adaptation motor. Work during the past decade has supported all three of these tenants.

To study the mechanical basis of adaptation, Assad and Corey (33) systematically measured adaptation rates for positive and negative bundle deflections. Based on these measurements and the previously measured bundle stiffness they developed a quantitative model that described the active, force-producing properties of the adaptation motor. They found that reducing calcium entry (by depolarizing the hair cell) reduced both climbing and slipping rates, but reduced slipping more. The difference in rates allowed for a simple test of the model: it predicted that depolarization would cause a decrease in calcium entry, which would allow the adaptation motors to climb up the stereocilia and thus increase tension on the channels. The added tension was expected to increase the channel open probability to ≈80% and therefore shift the I(X) curve by ≈120 nm, both of which were confirmed experimentally. Based on the geometry of the bundle (see Fig. 1) an increase in tip-link tension was predicted to pull an unrestrained bundle in the negative direction. The expected position change with depolarization, of ≈100 nm, was in close agreement with that actually measured by using high-resolution video microscopy (33). A final validation of the model was derived from an experiment in which the bundle movement caused by depolarization was abolished when tip links were cut with BAPTA, suggesting that tip links convey the tension that produces the bundle movement (6). Thus, this model accurately predicted movement of the bundle and supported a mechanical basis for adaptation.

On the other hand, morphological correlates of adaptation have been more difficult to obtain. Video microscopy with a resolution of 30–40 nm showed no gross movement of the stereocilia or cuticular plate during adaptation, thus the focus was shifted to smaller-scale molecular rearrangements (34). Thus far, it has not been possible to resolve the predicted changes (10–20 nm) in the position of the electron-dense plaques. On the other hand, cutting tip links with BAPTA should relieve tension and allow motors to climb even in an undeflected bundle. Indeed, BAPTA treatment was followed by an upward movement of the plaques by 50–70 nm, as measured from transmission electron micrographs (34).

A growing body of evidence supports the involvement of myosin in hair cell adaptation. First, using isolated frog stereocilia with the membranes removed, Shepherd *et al.* (34) showed that chicken muscle myosin II could move along the actin cores at a rate of 1–2 $\mu\text{m}/\text{sec}$ toward tips—the same direction as that predicted for the adaptation motor. Second, blockers of the ATPase cycle such as ADP βS , which arrest myosin while strongly bound to actin, block adaptation when dialyzed into the hair-cell cytoplasm (35). Similarly, phosphate analogs that arrest myosin while weakly bound should inhibit myosin force production and apparently cause release of resting tension on channels (36).

The myosin gene superfamily contains dozens of members, organized into 15 classes, eight of which occur in vertebrates. Identification of hair cell myosins has followed two strategies. Gillespie *et al.* (37) used vanadate trapping of adenine nucleotides to identify three putative myosins within stereocilia, of molecular masses 120, 160, and 230 kDa. The 120-kDa protein was labeled with an antibody to a myosin type I β , and this antibody particularly labeled the tips of stereocilia. Solc *et al.* (38) used degenerate PCR to amplify fragments of most myosins expressed in the hair-cell epithelium and found 10 different myosins from six different classes (I, II, V, VI, VII, and X). Interestingly, two of these myosins (VI and VIIa) cause inherited deafness in mice and/or humans when mutated (39, 40). The myosin-I β was cloned in full, and an antibody was raised against the tail domain (38). An extensive antibody study found that myosins I β , VI, and VIIa all are expressed by the hair cells, and all are in the stereocilia as well as elsewhere in the cells (41). These have molecular masses of approximately 120, 160, and 230 kDa, respectively. Recently, a myosin XV also has been discovered in stereocilia (42). Of all these myosins, however, only myosin I β is concentrated in the tips of stereocilia (Fig. 5 A and B) (41). Because localization with light microscopy cannot determine the relation of a myosin to the tip links, and because myosins might naturally climb to the tips of stereocilia unless otherwise prevented, the location of myosin I β was determined more precisely with ImmunoGold electron microscopy (43, 44). Myosin I β immunoreactivity was indeed associated with both end of the tip links, where it may link the channels to the actin cores (Fig. 5 C-E) (44).

Like most myosins, myosin I β has binding sites for regulatory light chains such as calmodulin. Three calmodulin molecules bind to myosin I β and confer a calcium dependence to the myosin activity (45, 46). Calmodulin is in stereocilia, especially concentrated at the tips (3, 47), and antagonists of calmodulin block adaptation* (49). It may be that calmodulin mediates the Ca²⁺ sensitivity of adaptation (49, 50).

Ca²⁺-Dependent Closure Model. The rate of adaptation depends on the Ca²⁺ concentration inside the tips of stereocilia. With high extracellular Ca²⁺, and low or slow internal calcium buffer, adaptation can become quite fast, with a time constant as short as 0.3 msec (16). This speed is probably too fast to be mediated by a tension adjustment system that requires an ATPase cycle of myosin. For instance, the calculated 1- to 2- $\mu\text{m}/\text{s}$ climbing rate of the adaptation motor and a myosin step size of 8 nm suggest an ATPase cycle time of 4–8 ms. A fast time constant is still compatible with a myosin motor if Ca²⁺ rapidly causes the myosin to release tension, for instance if Ca²⁺ reversibly causes the lever domain of myosin to become more compliant (50). However, such a mechanism could account only for a limited shift of the I(X) curve—perhaps 0.1 μm given the myosin lever dimensions and the bundle geometry—and fast adaptation can cause more shift than that.

To explain a rapid “twitch” in bundle movement, Howard and Hudspeth (11) proposed that Ca²⁺ binding directly to an intracellular site on the channel could shift the energetics of channel opening, such that the channel with calcium bound requires more force to reach the same open probability. That is, calcium entry would tend to close the channel (Fig. 3B). In an extension of this model, in which the energy of one of the two closed states is affected by Ca²⁺, Crawford *et al.* (29) suggested that Ca²⁺ binding is responsible for the adaptive shift of the I(X) curve. Although their initial model cannot explain adaptation to large deflections (33), a later model, in which calcium changes the “set point” of the channel, accounts for adaptation of up to 0.8 μm (32).

These two models—one in which Ca²⁺ relaxes tension and allows channel closure, and the other in which Ca²⁺ causes the channel to shut even with the same tension—have quite different predictions for the mechanical behavior of a hair bundle during

*Corey, D.P., Smith, W.J., Barres, B.A. & Koroshetz, W. J. (1987). *Soc. Neurosci. Abstr.* **13**, 538 (abstr.).

adaptation. In the motor model, a force that opens channels would quickly move the bundle a certain amount determined by the force and the stiffness of the bundle. Bundle stiffness is determined by two springs in parallel: the stereocilia pivot stiffness and the gating spring stiffness. If myosin motors slip to allow channel closure, the gating spring relaxes, which could be seen as a decrease in stiffness (chord stiffness but not slope stiffness). The bundle would move forward with the time course of adaptation. On the other hand, if adaptation happens because Ca^{2+} causes the channel to shut, the small movement of the channel's gate slamming shut would tighten the gating spring (Fig. 3E), in essence increasing the stiffness and pulling the whole bundle backward.

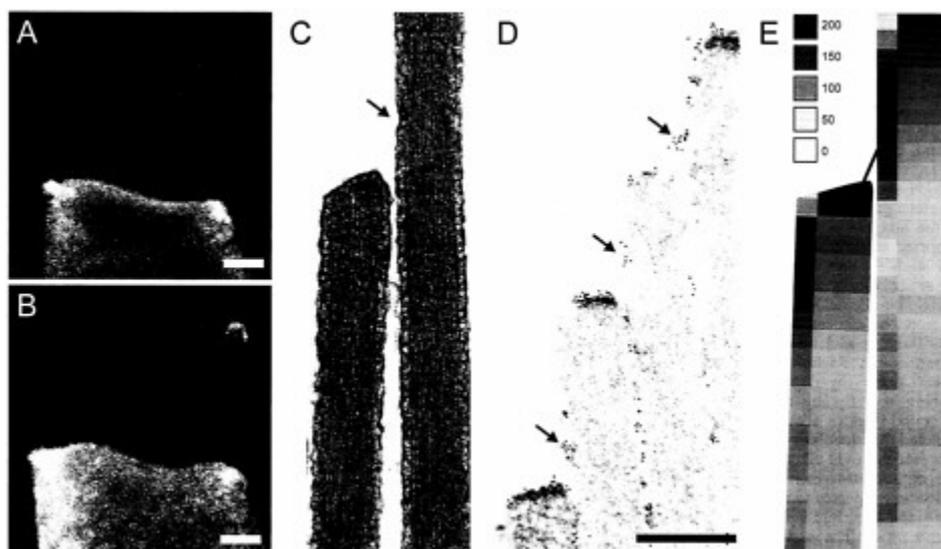


Fig. 5. Myosin $I\beta$ immunoreactivity in bullfrog hair cells. (A and B) Immunofluorescence of single dissociated cells shows myosin $I\beta$ in the cell body and also in the hair bundle where it is especially concentrated at the tips of stereocilia. The kinocilium also shows fluorescent label. (Scale bar=2 μm .) (C) Transmission electron micrograph ($\times 30,000$) showing a longitudinal section of two adjacent stereocilia. The arrow indicates the tip-link insertion and the electron-dense plaque, the presumed site of the adaptation motors. (D) Immunogold electron microscopy localizing myosin $I\beta$ at both ends of tip links. (Scale bar=500 nm.) (E) Summary of myosin $I\beta$ distribution in the tips of stereocilia. Gold particles marking the antibody were counted and averaged over many stereocilia in six bundles, and their density is indicated as particles/ μm^2 of surface membrane. The greatest density is within ≈ 200 nm of either end of the tip link. [Reprinted with permission from ref. 44 (Copyright 1998, Society for Neuroscience).]

Which behavior is seen with adaptation? In fact, both positive and negative going movements have been observed, with different time courses (refs. 11, 31, 51, and 52; R.Fettiplace, personal communication). Fig. 5A shows a rapid forward movement (arrowhead), followed by a quick hook back—the twitch—and then a slower and larger forward relaxation. Both the twitch and the slower relaxation depend on Ca^{2+} , becoming slower and smaller when Ca^{2+} is reduced. The termination of force allows the bundle to move back, but no twitch is seen on the return, at least after large deflections. On the other hand, a negative force evokes a slow negative relaxation of the bundle, and then a twitch is seen when termination of force allows the bundle to move forward toward the rest position.

These two mechanical behaviors, which are separable by time course and stimulus polarity, suggest that both proposed mechanisms of adaptation may be occurring in hair cells. The calcium closure might be fast (0.3–5 msec), and cause the quick twitch; a myosin slipping would be slower (10–100 msec) and may cause the slow relaxation. A careful analysis of the time course of adaptation in turtle hair cells has found that two time constants are needed to fit the transduction current. Moreover, they depend differently on external Ca^{2+} and internal buffer concentration, further indicating their separability (32). Comparison with a model of Ca^{2+} diffusion in stereocilia suggests that the fast phase is controlled at a Ca^{2+} binding site that is 20–50 nm from the site of Ca^{2+} entry and may be the channel itself (32). The slow phase is controlled at a more distant site, 150–200 nm from the site of entry, which corresponds well with the location of myosin $I\beta$. Wu *et al.* (32) did not speculate about molecular mechanisms of the two components, but it will be interesting to find out which parts of the transduction complex correspond to each phase.

Future Directions

How can we resolve the relative contribution of these two mechanisms to transducer adaptation in hair cells? Both the channel and the motor have measurable mechanical correlates of their activity, which can be recorded together with the receptor current. High-speed and high-resolution mechanical measurements, with concurrent voltage-clamp recording, could better clarify which way the bundle moves when fast or slow adaptation occur.

Ultimately, we need to test individual proteins for involvement in adaptation. At present, the only putative components of the transduction apparatus that have been cloned are myosin $I\beta$ and calmodulin—a remarkably meager portfolio in comparison to, say, phototransduction. Cloning the transduction channel would provide an important tool to study the calcium-binding form of adaptation. Even with candidates in hand, testing function is

problematic for components of the hair-cell transduction assembly: unlike voltage-gated ion channels, none of the proteins involved in transduction or adaptation are likely to display interesting behavior outside of their native environment. Instead, they must be disrupted in the intact hair cell and assayed physiologically to understand their function.

Although costly and time consuming, targeted gene deletion is one approach. Unfortunately, attempts to knock out myosin *I β* were found to be embryonic lethal (P.G. Gillespie, personal communication). To circumvent developmental complications one strategy would be to use hair cell-specific promoters, such as those for the $\alpha 9$ acetylcholine receptor (53) and Math1 (54), to drive expression of exogenous gene constructs in hair cells only. An alternative approach is to introduce dominant-negative fragments of candidate genes into hair cells by using adenoviral vectors (55). This approach offers several advantages: disruption of a single gene product in the fully developed system; ease of viral construction and delivery; identification of infected cells using green fluorescent protein; and the ability to study negative controls (uninfected cells) in the same tissue.

Recently the possibility of engineering drug sensitivity into endogenous proteins has been raised. Gillespie *et al.* (48) have designed a myosin *I β* construct that has a point mutation (Y61G) in its ATP binding pocket. Assays for ATPase activity and *in vitro* motility demonstrated that this construct behaved normally in the presence of dinucleotides and trinucleotides. However, introduction of bulky dinucleotide analogs arrested the activity of mutant, but not wild-type, myosin *I β* . Whether this construct alters transducer adaptation in sensory hair cells remains to be seen.

1. Flock, A., Cheung, H.C., Flock, B. & Utter, G. (1981) *J. Neurocytol.* **10**, 133–147.
2. Tilney, L.G., Derosier, D.J. & Mulroy, M.J. (1980) *J. Cell Biol.* **86**, 244–259.
3. Shepherd, G.M.G., Barres, B.A. & Corey, D.P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4973–4977.
4. Crawford, A.C. & Fettiplace, R. (1985) *J. Physiol.* **364**, 359–379.
5. Pickles, J.O., Comis, S.D. & Osborne, M.P. (1984) *Hear. Res.* **15**, 103–112.
6. Assad, J.A., Shepherd, G.M. & Corey, D.P. (1991) *Neuron* **7**, 985–994.
7. Jacobs, R.A. & Hudspeth, A.J. (1990) *Cold Spring Harbor Symp. Quant. Biol.* **55**, 547–561.
8. Corey, D.P. & Hudspeth, A.J. (1979) *Nature (London)* **281**, 675–677.
9. Corey, D.P. & Hudspeth, A.J. (1983) *J. Neurosci.* **3**, 962–976.
10. Corey, D.P. & Hudspeth, A.J. (1983) *J. Neurosci.* **3**, 942–961.
11. Howard, J. & Hudspeth, A.J. (1988) *Neuron* **1**, 189–199.
12. Zhao, Y., Yamoah, E.N. & Gillespie, P.G. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 15469–15474.
13. Kachar, B., Parrakal, M., Kurc, M., Zhao, Y.-D. & Gillespie, P.G. (2000) *Proc. Natl. Acad. Sci. USA* **97**, in press.
14. Denk, W., Holt, J.R., Shepherd, G.M.G. & Corey, D.P. (1995) *Neuron* **15**, 1311–1321.
15. Lumpkin, E.A., Marquis, R.E. & Hudspeth, A.J. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 10997–11002.
16. Ricci, A.J. & Fettiplace, R. (1997) *J. Physiol. (London)* **501**, 111–124.
17. Lumpkin, E.A. & Hudspeth, A.J. (1998) *J. Neurosci.* **18**, 6300–6318.
18. Roberts, W.M. (1993) *Nature (London)* **363**, 74–76.
19. Roberts, W.M. (1994) *J. Neurosci.* **14**, 3246–3262.
20. Ricci, A.J., Wu, Y.C. & Fettiplace, R. (1998) *J. Neurosci.* **18**, 8261–8277.
21. Yamoah, E.N., Lumpkin, E.A., Dumont, R.A., Smith, P.J., Hudspeth, A.J. & Gillespie, P.G. (1998) *J. Neurosci.* **18**, 610–624.
22. Street, V.A., McKee-Johnson, J.W., Fonseca, R.C., Tempel, B.L. & Noben-Trauth, K. (1998) *Nat. Genet.* **19**, 390–394.
23. Kozel, P.J., Friedman, R.A., Erway, L.C., Yamoah, E.N., Liu, L.H., Riddle, T., Duffy, J.J., Doetschman, T., Miller, M.L., Cardell, E.L. & Shull, G.E. (1998) *J. Biol. Chem.* **273**, 18693–18696.
24. Eatock, R.A., Corey, D.P. & Hudspeth, A.J. (1987) *J. Neurosci.* **7**, 2821–2836.
25. Holt, J.R., Corey, D.P. & Eatock, R.A. (1997) *J. Neurosci.* **17**, 8739–8748.
26. Shepherd, G.M.G. & Corey, D.P. (1994) *J. Neurosci.* **14**, 6217–6229.
27. Hacohen, N., Assad, J.A., Smith, W.J. & Corey, D.P. (1989) *J. Neurosci.* **9**, 3988–3997.
28. Assad, J.A., Hacohen, N. & Corey, D.P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2918–2922.
29. Crawford, A.C., Evans, M.G. & Fettiplace, R. (1989) *J. Physiol. (London)* **419**, 405–434.
30. Ricci, A.J. & Fettiplace, R. (1998) *J. Physiol. (London)* **506**, 159–173.
31. Howard, J. & Hudspeth, A.J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3064–3068.
32. Wu, Y.C., Ricci, A.J. & Fettiplace, R. (1999) *J. Neurophysiol.* **82**, 2171–2181.
33. Assad, J.A. & Corey, D.P. (1992) *J. Neurosci.* **12**, 3291–3309.
34. Shepherd, G.M.G., Corey, D.P. & Block, S.M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8627–8631.
35. Gillespie, P.G. & Hudspeth, A.J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2710–2714.
36. Yamoah, E.N. & Gillespie, P.G. (1996) *Neuron* **17**, 523–533.
37. Gillespie, P.G., Wagner, M.C. & Hudspeth, A.J. (1993) *Neuron* **11**, 581–594.
38. Sole, C.K., Derfler, B.H., Duyk, G.M. & Corey, D.P. (1994) *Aud. Neurosci.* **1**, 63–75.
39. Avraham, K.B., Hasson, T., Steel, K.P., Kingsley, D.M., Russell, L.B., Mooseker, M.S., Copeland, N.G. & Jenkins, N.A. (1995) *Nat. Genet.* **11**, 369–375.
40. Gibson, F., Walsh, J., Mburu, P., Varela, A., Brown, K.A., Antonio, M., Beisel, K.W., Steel, K.P. & Brown, S.D.M. (1995) *Nature (London)* **374**, 62–64.
41. Hasson, T., Gillespie, P.G., Garcia, J.A., MacDonald, R.B., Zhao, Y., Yee, A.G., Mooseker, M.S. & Corey, D.P. (1997) *J. Cell Biol.* **137**, 1287–1307.
42. Liang, Y., Wang, A., Belyantseva, I.A., Anderson, D.W., Probst, F.J., Barber, T.D., Miller, W., Touchman, J.W., Jin, L., Sullivan, S.L., *et al.* (1999) *Genomics* **61**, 243–258.
43. Steyger, P.S., Gillespie, P.G. & Baird, R.A. (1998) *J. Neurosci.* **18**, 4603–4615.
44. Garcia, J.A., Yee, A.G., Gillespie, P.G. & Corey, D.P. (1998) *J. Neurosci.* **18**, 8637–8647.
45. Zhu, T., Sata, M. & Ikebe, M. (1996) *Biochemistry* **35**, 513–522.
46. Burlacu, S., Tap, W.D., Lumpkin, E.A. & Hudspeth, A.J. (1997) *Biophys. J.* **72**, 263–271.
47. Walker, R.G., Hudspeth, A.J. & Gillespie, P.G. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2807–2811.
48. Gillespie, P.G., Gillespie, S.K., Mercer, J.A., Shah, K. & Shokat, K.M. (1999) *J. Biol. Chem.* **274**, 31373–31381.
49. Walker, R.G. & Hudspeth, A.J. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 2203–2207.
50. Gillespie, P.G. & Corey, D.P. (1997) *Neuron* **19**, 955–958.
51. Jaramillo, F. & Hudspeth, A.J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1330–1334.
52. Benser, M.E., Marquis, R.E. & Hudspeth, A.J. (1996) *J. Neurosci.* **16**, 5629–5643.
53. Zuo, J., Treadaway, J., Buckner, T.W. & Fritzsche, B. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 14100–14105.
54. Bermingham, N.A., Hassan, B.A., Price, S.D., Vollrath, M.A., Ben-Arie, N., Eatock, R.A., Bellen, H.J., Lysakowski, A. & Zoghbi, H.Y. (1999) *Science* **284**, 1837–1841.
55. Holt, J.R., Johns, D.C., Wang, S., Chen, Z.Y., Dunn, R.J., Marban, E. & Corey, D. P. (1999) *J. Neurophysiol.* **81**, 1881–1888.

COCHLEAR MECHANISMS FROM A PHYLOGENETIC VIEWPOINT

Geoffrey A. Manley*

Institut für Zoologie, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

The hearing organ of the inner ear was the last of the paired sense organs of amniotes to undergo formative evolution. As a mechanical sensory organ, the inner-ear hearing organ's function depends highly on its physical structure. Comparative studies suggest that the hearing organ of the earliest amniote vertebrates was small and simple, but possessed hair cells with a cochlear amplifier mechanism, electrical frequency tuning, and incipient micromechanical tuning. The separation of the different groups of amniotes from the stem reptiles occurred relatively early, with the ancestors of the mammals branching off first, approximately 320 million years ago. The evolution of the hearing organ in the three major lines of the descendants of the stem reptiles (e.g., mammals, birds-crocodiles, and lizards-snakes) thus occurred independently over long periods of time. Dramatic and parallel improvements in the middle ear initiated papillar elongation in all lineages, accompanied by increased numbers of sensory cells with enhanced micromechanical tuning and group-specific hair-cell specializations that resulted in unique morphological configurations. This review aims not only to compare structure and function across classification boundaries (the comparative approach), but also to assess how and to what extent fundamental mechanisms were influenced by selection pressures in times past (the phylogenetic viewpoint).

The hearing organs of modern amniotes (reptiles, birds, and mammals) show an almost bewildering variety of morphologies. This variety provides the functional morphologist with an exciting natural experiment to assess the functional consequences of structural diversity in an organ whose function largely depends on just this structure. To do this presupposes an understanding of (i) the evolutionary history of the amniotes, (ii) the extent and nature of the morphological variation, and (iii) those mechanisms of hearing that are directly influenced by morphology. This review emphasizes cochlear mechanisms that may have been influenced by morphological changes during amniote phylogeny. As is now customary, the term cochlear will be used loosely for the mechanisms involved in the hearing organ of the cochlear duct (Scala media) of all amniotes.

The present discussion is based on certain assumptions, the most important being that the hearing organs of all amniotes are homologous (1). They have a common ancestry, share a common structure, and develop from the same genetic substrate, and their position in the organisms' *Bauplan* is the same. Thus the functional units of the hearing process—the hair cells and their innervating nerve fibers—are also homologous.

A second assumption, based on the comparative anatomy and physiology of putatively primitive amniote hearing organs, is that the archetypal auditory papilla was short (≈ 1 mm) with only a few hundred hair cells. It is assumed that these hair cells (i) had inherited electrical tuning from their vestibular-system ancestors, (ii) were innervated by both afferent and efferent nerve fibers, and (iii) contained an active process in their stereovillar bundles (2).

The final assumption is that evolution is a conservative process. Rather than developing functions or structures *de novo*, existing structures and processes tend to be modified, sometimes to accomplish new tasks. This conservatism also applies, of course, to subcellular molecular mechanisms. Using comparative studies of extant vertebrates, I will estimate how inner-ear mechanisms have been modified during phylogeny. To begin, it is appropriate to briefly introduce middle-ear evolution and the different hearing-organ morphologies of modern amniote lineages.

The Tympanic Middle Ear as the Initiator of Profound Change

The development of a tympanic middle ear in the late Paleozoic was crucial to the initiation of selection pressures for further phylogeny of the hearing organ. In early stem reptiles, there was little resembling what we would now refer to as a tympanic middle ear, i.e., an efficient impedance-matching device between air and water, with a tympanum moved by sound-pressure variations in air. Recent paleontological studies indicate that a tympanic ear was established in all lineages during the Triassic [250–220 million years (MY) ago, Fig. 1], 150 MY after the origin of the amniotes (3). This was 100 MY after mammal-like reptiles (synapsids) and 75 MY after the avian-crocodilian and lizard-snake lineages (referred to below as avian and lizard lineages) diverged from stem reptiles. We do not know which selective pressures led to the essentially simultaneous development of tympanic ears in all lineages. The new middle ears enabled the perception of higher frequencies, producing selection pressures for sensitivity to high-frequency sounds (>1 kHz).

This transformation had very profound consequences for the auditory papilla, for which the following sequence of evolutionary developments is likely. The refinement of micromechanical tuning extended the frequency range and made papillar elongation an advantage, enlarging the frequency space constants (Fig. 1). The increase in the number of sensory cells permitted the specialization of two hair-cell populations. In lizards, the two populations are at different locations along the tonotopic gradient, whereas in birds and mammals, they are separated across the papilla but present all along the tonotopic gradient. In birds and mammals, a new division of labor specialized one of the two populations for cochlear amplification. The independent evolution of larger and more specialized auditory organs also led to size increases and specialization of the neural auditory pathway. Their parallel evolution may explain the difficulties of establishing homologies between brain nuclei in different groups (e.g., ref. 4).

The tympanic ear of mammals evolved differently to that of other groups. In the mammalian lineage, the jaw joint shifted from the primary to the more superficial secondary joint. The quadrate (malleus) and articular (incus), in a quirk of phylogeny, were integrated into the chain of three ossicles. This initially achieved the same result as the single-ossicle system of the avian

*E-mail: geoffrey.manley@bio.tum.de.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: BM, basilar membrane; IHC, inner hair cell; OAE, otoacoustic emissions; OHC, outer hair cell; SHC, short hair cell; THC, tall hair cell; TM, tectorial membrane.

and lizard lineages. Once selection pressures for responses to higher frequencies had begun to work, however, the mammalian solution was, quite accidentally, better at transmitting high frequencies (Fig. 2) (5).

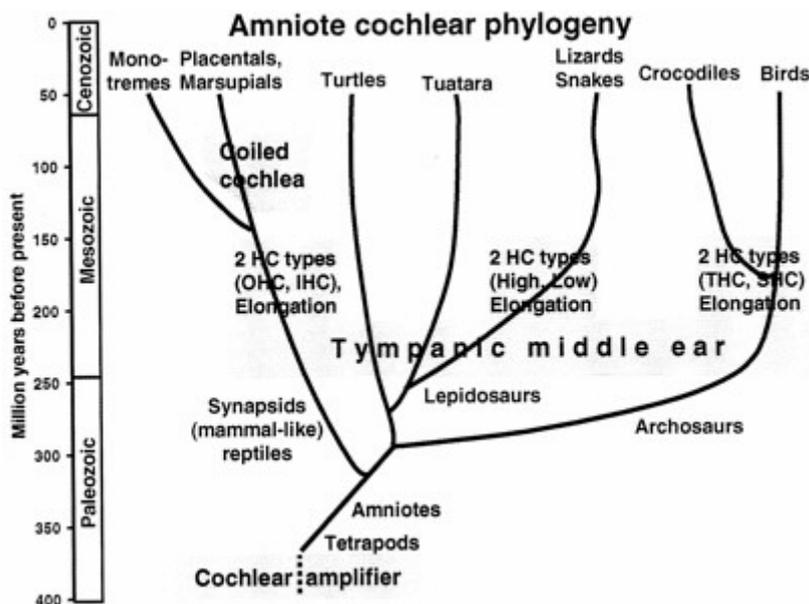


Fig. 1. Highly schematic representation of the amniote phylogenetic tree over 400 million years to illustrate the approximate time of origin of particular features of auditory systems. Amniotes arose from the earliest tetrapods early in the paleozoic and inherited from them a simple hearing organ with a cochlear amplifier in the stereovillar bundles. Apart from the lineages to the turtles and the Tuatara, that remained primitive in a number of respects, three main lineages to modern amniotes are distinguished here. Splitting off first were mammalian ancestors, which gave rise to both the egg-laying monotremes and the marsupial-placental line. Later, the archosaur line originated and led to the dominant land organisms of the mesozoic. Of these, only the crocodile-alligator and bird groups survived to modern times. The last group to split off was the lizards and snakes within the lepidosaurs. The tympanic middle ear originated independently in all groups during the Triassic, initiating the evolution of unique configurations of papillae, with all groups showing papillar elongation and hair-cell specializations. In mammals IHC and OHC, in birds THC and SHC populations developed. In lizards, the populations segregated along frequency lines (low- and high-frequency populations). Because the hair-cell populations in the monotreme and marsupial-placental mammal groups are so similar, they almost certainly arose before these lineages diverged. The same applies to the birds and Crocodylia. In lizards, there are great family-specific variations, suggesting that these hair-cell populations arose soon after the Triassic. Because monotremes do not have a coiled cochlea, coiling almost certainly developed in the marsupial-placental lineage. [Modified after ref. 5 and used with permission of Wiley-VCH Press (Copyright 2000, Wiley).]

The Lineages of Modern Amniotes and Their Characteristic Hearing-Organ Morphologies

There are four basic types of amniote ear, in largely natural, i.e., phylogenetic, groups (9). I will refer to mammalian, avian (including *Crocodylia*), lizard (including snakes), and turtle (standing for a stem-reptile or unspecialized type) lineages. Mammalian, avian, and lizard papillae show unique constellations of features, suggesting that these features developed early in their respective phylogenies. Except for the turtles, all groups developed a sensitivity to high frequencies through unique morphological changes (2).

Turtles and Tuataras Probably Represent the Unspecialized State of the Ear. Turtles (and the Tuatara lizard, *Sphenodon*) are regarded as primitive (9) and have the least specialized hearing organ (Fig. 3) (10–12), most likely to resemble that of early stem reptiles. The hair cells are unspecialized, innervated by both afferent and efferent nerve fibers, and respond only to low frequencies (<1 kHz). The stereovillar bundles of almost all hair cells are uniformly oriented with weak morphological gradients that play only a minor role in determining hair-cell response frequencies. Instead, the ion-channel complement of hair-cell membranes create electrical resonances at preferred frequencies (e.g., ref. 13). The calibration of the individual cell is achieved by varying the number and kinetics of the channels (Fig. 4D), with preferred frequencies from <100 Hz apically to about 600 Hz basally.

The basilar papilla was the first hair-cell organ to form over a moveable membrane (basilar membrane, BM). In turtles, as in lizards (see below), the BM shows no special frequency selectivity (14). In birds, there is some selectivity (15). In mammals, BM selectivity is the same as that of primary auditory afferents (e.g., ref. 16), but much is because of the motor activity of hair cells; the passive BM response is only crudely frequency selective. The primary focus in understanding the function of the inner ear needs to be on the hair cells themselves.

Lizards—A Playground of Evolution. The morphology of lizard auditory epithelia varies widely, but tends to be consistent within each family (10, 12, 17). It varies in length from <100 μm to >2 mm and in the number of hair cells from <60 to >2,000 (Fig. 5). Hair cells may be covered by a continuous tectorial membrane (TM), by a TM that is divided into a chain of sallets, or have no TM at all (Fig. 5). The tonotopic organization (the arrangement

of frequencies along the epithelium) may run from apex to base or from base to apex. Each lizard family shows particular combinations of features, and hearing-organ structure can even be species-specific.

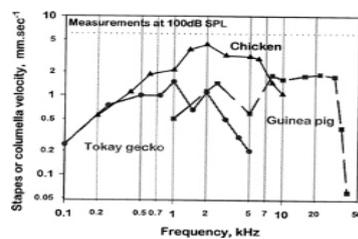


Fig. 2. Middle-ear sensitivity in representatives of three amniote groups, shown as the velocity of the center of the eardrum (the tip of the malleus or extracolumella) as a function of frequency at 100 dB sound pressure level (SPL). At this sound pressure, air particles have a velocity of 6 mm/sec (dashed line). The Tokay gecko (gray line, ●) is a sensitive lizard (6), the chicken (black line, ▲) represents birds (7), and the guinea pig (gray line, ■) represents mammals (8). The main difference observed is not in sensitivity, but in the upper frequency limits and the high-frequency flanks.

To reconstruct the phylogeny of the lizard hearing organ, it is necessary to recognize common features between lizard families. The first is that lizard papillae always show two hair-cell types (Fig. 6) (10). One papillar area contains hair cells with a greater basal diameter, large numbers of, and larger, afferent nerve fibers and an efferent innervation. In most lizard papillae, hair-cell bundles in such areas all have the same (abneural) orientation. Functionally, these cells always respond to low frequencies (below about 1 kHz; ref. 17). This area is like the entire papilla of turtles, with weak morphological gradients and a similar upper limit (<1 kHz). Its frequency selectivity also may be largely determined by electrical tuning, but there is only indirect evidence for this (18).

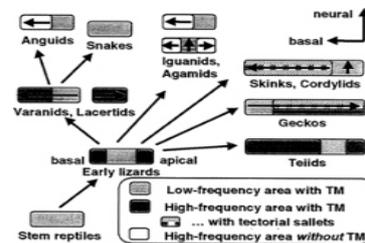


Fig. 3. A schematic summary of the structure of the auditory papilla in a primitive amniote, the red-eared turtle. Most of the hair cells are placed over the BM and covered by a thick TM (yellow). There is only one type of hair cell, which is innervated by both afferent and efferent fibers (*Right*).

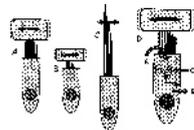


Fig. 4. (A-C) Basic hair-cell structures contributing to micromechanical frequency tuning. In most cases, a tectorial structure covers the hair-cell bundle. A hair cell with a taller stereociliary bundle and larger tectorial cover (A) responds best to much lower frequencies than a hair cell with a shorter bundle and less massive tectorial structure (B). Equivalent frequencies in hair cells without tectorial structures (C) result from much taller bundles. (D) Schematic of the ion channels involved in electrical tuning of hair cells. The flow of ions (mainly K^+) through transduction channels depolarizes the cell and activates voltage-sensitive Ca^{2+} channels in the baso-lateral membrane, raising the Ca^{2+} concentration [Ca^{2+}] in the cell and activating Ca^{2+} -sensitive K^+ channels. This leads to K^+ outflow, hyperpolarization, and a new cycle. Channel number, kinetics, and the temperature determine the frequency of oscillation. In turtles, frequencies between 50 and 600 Hz have been measured, but in other amniotes, higher frequencies are reached (see text).

The second hair-cell type is characterized by its smaller size, by smaller and fewer afferents, and the complete lack of an efferent innervation (Fig. 6). Almost all such regions have

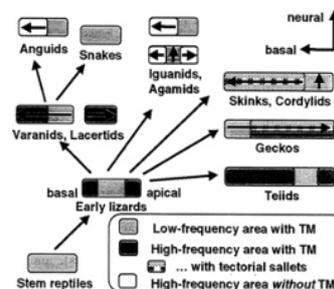


Fig. 5. Highly schematic representation of a possible evolutionary sequence of the papillae of modern lizard families. Where known, the direction of the tonotopicity, from low to high frequencies, is shown (arrow). From stem reptile papillae with uniform hair cells, a papilla arose with new, mirror-imaged, micromechanically tuned hair-cell areas at both ends, flanking the low-frequency (green) area. From this, the various papillar configurations of different lizard families and the snakes can be derived as shown, including the reversed tonotopic organization in geckos. Placing similarly formed papillae together does not necessarily imply close systematic relationships. In different families, the TM over the high-frequency areas was either retained (uniformly blue areas), divided up into sallets (patterned blue areas), or lost (yellow areas) (see text; partly after ref. 10). [Modified after ref. 5, and used with permission of Wiley-VCH Press (Copyright 2000, Wiley).]

groups of neurally and abneurally oriented hair-cell bundles (bidirectional orientation). These cells always respond to frequencies above about 1kHz, with an upper limit of at least 4kHz and are micromechanically tuned. The high-frequency areas show a large variation in the TM; some totally lack a TM. There are sometimes two high-frequency areas located apically and basally (Fig. 5) (10). If there is only one such area, it can be apical (e.g., geckos, in which case the tonotopic organization is reversed as compared with other amniotes, ref. 19) or basal (almost all others).

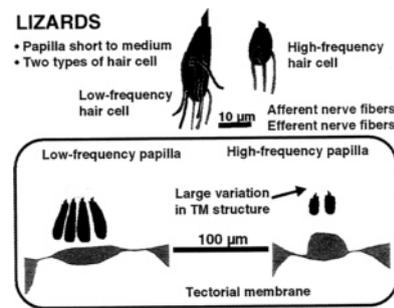


Fig. 6. A schematic summary of the structure of the auditory papilla of lizards, which always have a low- and a high-frequency area. Over the latter, the tectorial structure (yellow) is highly variable between and even within families and is missing in some groups (see Fig. 5). Low- and high-frequency hair cells differ both in their size and their innervation pattern; high-frequency hair cells (*Upper Right*) never receive an efferent innervation.

The evolution of the tympanic middle ear probably initiated the development of the high-frequency hair-cell areas of stem lizards; these are thus a synapomorphy (a derived common feature) of lizards. The two types of lizard hair cells are confined to separate frequency ranges and thus not functionally equivalent to the hair-cell populations of mammals or birds.

As in turtles, the lizard BM is not involved in frequency selectivity (20, 21), and the frequency selectivity of afferent nerve fibers is much higher than that of the BM (20). The primary stimulatory motion of hair cell bundles is transverse to the BM (e.g., ref. 22), and frequency selectivity depends on bundle and TM micromechanics (23).

Birds and Mammals—The Parallel Evolution of Specialized Hair-Cell Types. Avian and mammalian auditory papillae are characteristically different and easily recognizable. Both have specialized hair-cell populations located across the width of the papilla, essentially at all frequency locations (24), and thus within a continuous tonotopic organization (25). Both groups have— independently— developed responses to high frequencies, in some birds up to 10 kHz, in some mammals even beyond 100 kHz (26, 27). Papillar elongation was generally much more extensive than the maximum of 2 mm found in lizards. Some owl papillae reach 11 mm and some whale papillae 105 mm. The coiled cochlea, which evolved after the divergence of the marsupial-placental line in the late Mesozoic (Fig. 1), was probably simply a mechanism for accommodating a long papilla.

Although the anatomical differentiation of different hair-cell types in birds is not as clear-cut as in mammals, important structural parallels exist between the tall and short hair cells (THCs and SHCs) of birds and the inner and outer hair cells (IHCs and OHCs) of mammals, respectively (Fig. 7) (24). THCs and IHCs are the less specialized hair cells and receive a strong afferent innervation. OHCs are innervated nonexclusively by relatively few afferent fibers ($\approx 5\%$), and SHCs receive no afferent innervation at all (28).

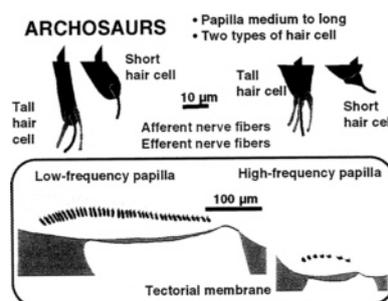


Fig. 7. A schematic summary of the structure of the auditory papilla of birds, which are between 2 and 11 mm in length and contain between 3,000 and 17,000 hair cells. The hair cells form a continuum, with the tallest cells at the apical end (*Left*, transverse section and typical hair-cell shape), where there are many hair cells across the papilla. Hair cells become shorter toward the base, where the number of hair cells across the papilla is much smaller (*Right*, transverse section and typical hair-cell shapes), and toward the abneural side. All hair cells are covered by a wedge-shaped TM (yellow). Short hair cells have no afferent innervation.

In mammals, the transition from IHC to OHC is sudden, and the populations are separated by specialized supporting cells. In birds, the transition is usually gradual (Fig. 7). Phylogenetically, the most interesting feature of avian and mammalian papillae that accompanied the development of high-frequency hearing is the setting aside of hair cells as effectors in a cochlear amplification process, as follows: IHCs and THCs retain the classical sensory function, passing information via afferents to the brain. OHCs and SHCs, in contrast, respond to low-level stimuli by creating motion that amplifies stimuli, that are passed on to IHCs or THCs (2). This occurred despite the fact that birds have retained electrical tuning of hair cells (25) and married it seamlessly to micromechanical tuning, whereas mammals appear to have abandoned the ancestral electrical tuning and to rely completely on micromechanics. Perhaps the latter was possible because of—or even necessary for—the development of a new cellular mechanism of amplification (see below).

A uniquely evolved feature of mammals is the intimate involvement of the BM in the response, such that its frequency selectivity is essentially identical to that of IHCs and afferent fibers (e.g., ref. 16). The response is thus a feature of the entire organ, and no component is separable without reducing sensitivity and selectivity. This linking is not so well developed in birds (15, 29), perhaps because those hair cells that connect to most of the afferent fibers are not over the free BM, but over the limbus.

Frequency-Selectivity Mechanisms and Tonotopicity from a Phylogenetic Viewpoint

All vertebrate hearing organs, including those of fish and amphibians, consist of frequency-selective hair cells. Of the two mechanisms of frequency selectivity in hair cells (30), the phylogenetically older is electrical tuning of the hair-cell membrane, as described above for the turtle inner ear (Fig. 4). Strong evidence for such tuning is available for hair cells of the frog sacculus and the avian basilar papilla. At the higher tempera

tures of endotherms such as birds, electrical tuning should function to at least 4 kHz (31). Preferred intervals in the spontaneous activity of primary auditory neurons, a probable correlate of electrical tuning, have been reported from turtles, lizards, and birds (25). In the barn owl, Köppl (32) found preferred intervals in fibers of characteristic frequencies up to 5 kHz. Because both electrical tuning and phase locking work well at low frequencies, there was presumably little selective pressure for developing new selectivity mechanisms until the detection of high frequencies became important.

Responses to higher frequencies are best mediated by the second mechanism providing frequency selectivity, micromechanical tuning (Fig. 4 A-C). Inner-ear structures are extremely small and have high resonance frequencies. The response frequency is determined by the mass and stiffness of their stereovillar bundles and the mass of any tectorial material covering the hair cells or of fluid coupled to the hair cells. In the later evolution of some groups, especially mammals, larger structures such as the BM were recruited into oscillating units. This resulted in significant changes in supporting-cell structure, influencing the mechanical properties of the entire organ. Thus, although micromechanical (at the hair-cell level) and macromechanical (involving larger structural units) tuning often are distinguished for discussion purposes, in the living organism they form a continuum. Phylogenetic variations on micromechanical tuning are discussed below.

The Factors Influencing Micromechanical Tuning Illustrated by Using Lizard Papillae. Almost the entire variability in papillar structure between different lizard families is in the high-frequency, micromechanically tuned regions (6). Thus phylogenetically, high-frequency hearing arose near the time when the lizard families diverged. Beginning with small, micromechanically tuned areas at both ends of the low-frequency area, each family specialized these high-frequency areas differently. Despite this, the upper frequency limit is similar in all cases and lies between about 4 and 7 kHz (17).

The two most important determinants of frequency selectivity of micromechanically tuned units in lizard papillae are (i) the length of the papilla (correlated with the number of hair cells), and (ii) the presence or absence of a tectorial structure and how that structure is shaped (Fig. 8). In a model of frequency tuning developed for the Tokay gecko (23), it was shown that tectorial structures increase sensitivity (the amplitude factor for free-standing hair bundles is only one-third of that for a tectorial sallet system) and frequency selectivity of hair cells (the amplitude maxima of hair-cell bundles without a TM are significantly less sharp than those with one; see also Fig. 8). The difference is largest when comparing free-standing hair cells with those covered by a chain of sallets. Under a continuous TM, the coupling between neighboring frequency regions is stronger, and in short papillae this would negatively influence frequency selectivity. The presence of essentially independent tectorial sallets in hair-cell areas in geckos, skinks, and cordylid lizards thus can be regarded as a method of micromechanical optimization in a small system covering a wide frequency range.

Most lizards with small basilar papillae (e.g., iguanids, anguids, and agamids) have only free-standing hair cells in the high-frequency papillar area. Without a TM, such hair cells can respond at the lowest to about 1 kHz (Fig. 4C); even then, their bundles contain few stereovilli, and these are remarkably tall (>30 μm ; Fig. 4C) (6). Coupling the small number of hair cells by using tectorial material would allow only very few distinct frequency responses. Loss of the TM in very small papillae was thus probably a phylogenetic mechanism for deriving the maximum amount of information on frequencies. The cost involved was some loss of sensitivity and frequency selectivity (Fig. 8) (6).

Hair-cell systems with continuous TMs tend to be large, thus neighboring hair cells are tuned to similar frequencies. Any minor loss of frequency selectivity caused by TM coupling is more than matched by the resulting gain in sensitivity. The larger the papilla, the less detrimental to frequency selectivity is the coupling of hair cells and adding a TM is advantageous. All large lizard papillae and all avian and mammalian papillae have hair cells covered by a TM. In intermediate cases, a sallet-chain TM optimizes the various influences on sensitivity and selectivity.

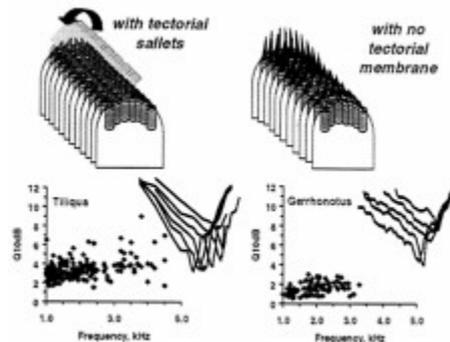


Fig. 8. The effect of TM coupling on the responses of lizard papillae. (Left) The salletal TM is shown lifted, to reveal hair-cell bundles. In the alligator lizard *Gerrhonotus* (Right), the high-frequency papillar area contains free-standing hair cells that lack a TM over their bundles. The requisite frequency range is attained by a strong gradient in bundle height along the papilla. Here, tuning sharpness of afferent fibers is poor (given as the $Q_{10\text{dB}}$ of tuning curves—the center frequency divided by the tuning bandwidth at 10dB above the best sensitivity). (Inset) Representative tuning curves are shown. Where hair cells are attached to tectorial sallets (Left, represented by the Australian Bobtail lizard *Tiliqua*), the coupling of neighboring hair cells leads to greater sensitivity and higher frequency selectivity, manifest as greater $Q_{10\text{dB}}$ values. In addition, the tuning curve flanks in *Tiliqua* are steeper (6).

Arranging Frequency Responses Systematically: The Phytogeny of Tonotopicity. Next to frequency selectivity, a tonotopic organization (i.e., the systematic arrangement of frequencies along the epithelium) is a primordial feature of hearing organs of terrestrial vertebrates. Hearing epithelia always present frequency as their main organizational feature, presumably because of developmental processes creating gradients along certain axes. In amniotes, the base of the hearing epithelium (nearest the sacculus) is almost always tuned to the highest frequencies. After a systematic drop in frequencies along the epithelium, the lowest frequencies are found at the apex. As so often, some lizards are an exception.

The ancestral lizard hearing epithelium was probably small, with the primordial low-frequency hair-cell area in the center, and incipient high-frequency areas at both ends (Fig. 5) (17). From this archetype, each lizard family or family group evolved its own arrangement (Fig. 5). In most iguanids, anguids, and agamids, the archetype was preserved with perhaps a reduction in size. The selection pressure for the preservation of redundant mirror-image hair-cell areas was presumably not very great, however, and some genera of these families lost the apical high-frequency area. The apical high-frequency area was also lost in most other lizard families, resulting in a tonotopic organization resembling that of birds and mammals (Fig. 5). That there is no universal law of tonotopic arrangements is shown by the geckos, which lost the basal high-frequency area of the archetype, and thus the tonotopic arrangement is the reverse of the amniote norm (Fig. 5). Other, more bizarre cases exist

among lizards, in which a physical separation developed between one of the high-frequency areas and the rest of the papilla, resulting not only in two subpapillae, but in a diversification of the frequencies in the two high-frequency areas (Fig. 5). This is found in the monitor-lizard family (Varanidae), the Lacertidae, and some isolated iguanid lizards (6).

In birds, the gradients for both electrical and micromechanical tuning run parallel and in register. Whether they develop simultaneously in ontogeny is unknown: Perhaps one gradient develops first and influences the expression of genes responsible for the other gradient. In fact, there are many gradients, both along and across the avian papilla, resulting in hair cells that are all unique. Every avian hair cell has a phenotype differing slightly from that of its neighbors (33).

Although the avian pattern is rather consistent across species, the barn owl has a severely distorted tonotopic organization. In its auditory fovea, the octave vital for prey detection (4–8 kHz) occupies half of the papillar length (e.g., ref. 34). Similarly, some mammalian papillae show auditory foveae; in some species of bats, space constants of more than 50 mm/octave occur in frequency ranges used in detecting prey. These foveae are thought to have developed independently in different bat groups, because their anatomical and physiological characteristics differ (35). The immense amount of afferent information from foveal regions is processed by enlarged regions of brain nuclei. The phylogeny of such foveae and the complexity of lizard tonotopicitities demonstrate the flexibility of the genetic substrate determining tonotopic arrangement, responding to selection pressures within a short evolutionary time.

▲ **Increasing Papillar Length During Evolution: Space Constants Do Matter.** One of the most consistent findings across the amniotes is an increase in length of the auditory epithelia during phylogeny. The increase in length is proportionately greater than the increase in the upper frequency, especially in birds and mammals. Thus a mammal with a papillar length of 25 mm and an upper limit of 50 kHz has a hearing range of about 10 octaves and a frequency space constant of 2.5 mm/octave (an average for mammals; Fig. 9). This compares to ≈ 3 octaves and a space constant of 0.3 mm in the putative ancestral form. In birds and in the high-frequency areas of large lizard papillae, the average space constant is < 1 mm/octave. In small lizard papillae, it can be < 0.1 mm. In some cases, only 20 columns of hair cells code for 2.5 octaves. But why are space constants important?

For electrically tuned hair cells, the frequency response of neighboring hair cells is, in principle, quite unimportant. The development of micromechanical tuning, however, placed a greater premium on space, because micromechanical influences from neighboring hair cells became critically important, as seen above. In the evolution of most amniote papillae, a continuous TM was kept, but the papilla was strongly elongated, increasing the number of hair cells and resulting in larger frequency space constants (Fig. 9). This maintained sensitivity and selectivity while simultaneously increasing parallel processing. Questions about space constants and those related to tonotopicity are thus interwoven.

Coding of Intensity from a Phylogenetic Viewpoint

The sound-driven activity of auditory afferents also encodes information on the signal level. Because the natural world offers a huge range of intensities, we might expect phylogenetic adaptations to cope with it. Range fractionation is one solution, in which individual nerve fibers have different thresholds and code over different level ranges. A wide threshold range across primary auditory nerve fibers is the rule (e.g., refs. 6, 25, and 37). In mammals, auditory afferents can be divided into three groups (that are, however, part of a continuum). The most sensitive fibers contact the abneural edge of IHCs (Fig. 10) and have a small dynamic range. The two other types of fiber are less sensitive, have wider dynamic ranges, and contact the neural side of IHCs (38). Yates *et al.* (40) showed that these types of rate-intensity function reflect the existence of a cochlear amplifier driving the BM, with a strong compressive nonlinearity

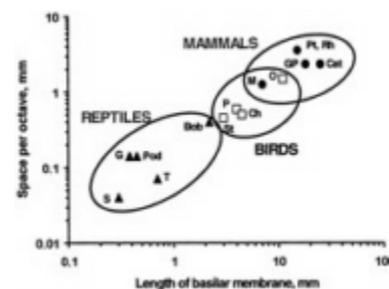


Fig. 9. The length of the BM and thus of the auditory organ correlates with the space constant for frequency (averaged along the papilla) in reptiles (\blacktriangle), birds (\square), and mammals (\bullet). Because their hearing range did not increase in proportion to the length of the cochlea, avian and mammalian papillae gained much more space for the analysis of any given octave. The increase was greater in most mammals than in most birds, but the distributions overlap. The resulting increase in hair cells and nerve fibers per octave increased parallel processing. Lizards: G, alligator lizard, *Gerrhonotus*; T, turtle; Pod, European lizards of the genus *Podarcis*; S, the granite spiny lizard *Sceloporus*; Bob, Australian Bobtail lizard *Tiliqua*. Birds: St, starling, *Sturnus*; Ch, chicken, *Gallus*; P, pigeon, *Columba*; O, barn owl, *Tyto*. Mammals: M, mouse, *Mus*; GP, guinea pig, *Cavia*; Cat, house cat, *Felis*; Pt, Rh, two bat species, *Pteronotus*, the mustached bat, and *Rhinolophus*, the horseshoe bat. For refs see also refs. 30, 34, and 36. [Modified after ref. 5, and used with permission of Wiley-VCH Press (Copyright 2000, Wiley).]

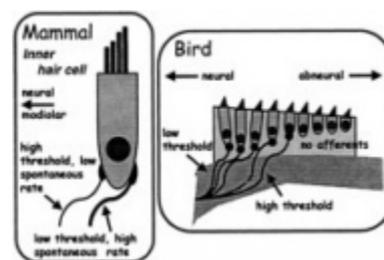


Fig. 10. Innervational patterns of primary afferents in mammals and birds that encode auditory sensitivity partly through the specialization of afferent fibers. In mammals (Left), the most sensitive fibers, which also have a higher spontaneous activity, are thicker and innervate the outer side of the IHC. Less sensitive fibers, which have lower spontaneous rates, innervate the modiolar side of the IHC (38). In the starling (Right), different sensitivities of afferent fibers contact different hair cells. The more sensitive fibers contact hair cells near the neural edge, less sensitive fibers hair cells near the middle of the papilla (e.g., ref. 39). Hair cells lying further abneurally have no afferent innervation (28). [Modified after ref. 5, and used with permission of Wiley-VCH Press (Copyright 2000, Wiley).]

above a certain level. Creating a more effective coding of sound intensity also may have been a selection pressure in the development of amplification systems.

In birds, primary afferent fibers often show a range of thresholds of >50 dB (6, 25). All three types of rate-intensity function described for mammals also exist, but in different proportions (40, 41). The particular pattern of rate-intensity types in birds suggests that the cochlear amplifier in birds acts locally rather than globally on the entire organ (40). There is evidence that the most sensitive afferents innervate hair cells near the neural edge of the papilla (Fig. 10) (25). In the emu, more than half the fibers code for a dynamic range greater than 50 dB (40).

The Phylogeny of the Cochlear Amplifier

Hair cells harbor a mechanism capable of increasing motion induced by low-level sounds (2, 42–44). Active motor mechanisms at the hair-cell level are responsible for the cochlear amplification at the organ level. In mammals, several phenomena correlate with the activity of a cochlear amplifier. These include the high sensitivity and frequency selectivity of hearing and their sensitivity to physiological insults, the existence of compressive nonlinearities in afferent fiber rate-level functions, and the existence and characteristics of otoacoustic emissions (OAE), especially spontaneous OAE (SOAE). SOAE are peaks in the sound-pressure spectra in the external meatus that are present in the absence of external stimuli and represent spontaneous hair-cell motion (45). All these phenomena also exist in nonmammals (2, 44). SOAE are more common in some nonmammals and resemble mammalian SOAE to a remarkably high degree. There is thus no reason to doubt the existence of a cochlear amplifier mechanism in nonmammals, including amphibians. The most recently discovered cochlear mechanism may thus be phylogenetically one of the oldest.

Active motion in hair-cell bundles has been demonstrated in vestibular hair cells (e.g., refs. 46–49) and auditory hair cells of the turtle (50). In nonmammals, it appears that the force of the cochlear amplifier operates parallel to the surface of the papilla, and thus at the interface between the TM and the hair cells. In birds, this motion may be transmitted via the TM to THCs, producing maximal motion amplitudes at a point near the neural edge of the epithelium (29, 51). This is where the starling, for example, has its most sensitive hair cells (Fig. 10) (39). Such motions might be only very poorly reflected in BM motion. In mammals, there is evidence that the OHCs have an amplification motor located in the lateral cell membrane that causes cell contraction and elongation in a plane almost at right angles to the BM and thus the movement is reflected in BM motion (52).

The two putative active processes correlate with morphological differences between birds and mammals. In birds, there are more stereovilli per hair cell than in mammals (25), which might be expected if the amplifier were located in the bundles. In contrast, mammals have a thin BM and papillae that are full of specialized supporting cells. Mammals may have abandoned the active motion of hair-cell bundles, at least at high frequencies, but there is as yet no evidence for this. In mammals as in nonmammals, a bundle motor also would be an effective way of coupling cellular power into organ motion (53).

In sum, it can be said that the most important changes in cochlear mechanisms during phylogeny were initiated by changes in the middle ear. This led in most lineages to a predominance on micromechanical tuning, a profound elongation of the papilla, and the specialization of hair cells that went as far as to generate a division of labor in birds and mammals. Many of the most fundamental cochlear mechanisms have, however, changed little or not at all during amniote phylogeny. Plus ça change, plus cés la même chose.

This work was supported by the Deutsche Forschungsgemeinschaft within the SFB 204 "Gehör" and DFG-Grant MA 871/10–2.

1. Nielsen, C. (1995) *Animal Evolution* (Oxford Univ. Press, Oxford).
2. Manley, G.A. & Köppl, C. (1998) *Curr. Opin. Neurobiol.* **8**, 468–474.
3. Clack, J.A. (1997) *Brain Behav. Evol.* **50**, 198–212.
4. Carr, C. (1992) in *The Evolutionary Biology of Hearing*, eds. Fay, R.R., Popper, A.N. & Webster, D.B. (Springer, New York), pp. 511–543.
5. Manley, G.A. (2000) in *Auditory Worlds: Sensory Analysis and Perception in Animals and Man*, eds. Manley, G.A., Fastl, H., Kössl, M., Oeckinghaus, H. & Klump, G.M. (Wiley, Weinheim), pp. 7–17.
6. Manley, G.A. (1990) *Peripheral Hearing Mechanisms in Reptiles and Birds* (Springer, Heidelberg).
7. Saunders, J.C. (1985) *Hear. Res.* **18**, 253–268.
8. Manley, G.A. & Johnstone, B.M. (1974) *J. Acoust. Soc. Am.* **56**, 571–576.
9. Carroll, R.L. (1988) *Vertebrate Paleontology and Evolution* (Freeman, New York).
10. Miller, M.R. (1992) in *The Evolutionary Biology of Hearing*, eds. Webster, D.B., Fay, R.R. & Popper, A.N. (Springer, New York), pp. 463–487.
11. Sneary, M.G. (1988) *J. Comp. Neurol.* **276**, 573–587.
12. Wever, E.G. (1978) *The Reptile Ear* (Princeton Univ. Press, Princeton, NJ).
13. Fettiplace, R. & Fuchs, P.A. (1999) *Annu. Rev. Physiol.* **61**, 809–834.
14. O'Neill, M.P. & Bearden, A. (1995) *Hear. Res.* **84**, 125–138.
15. Gummer, A.W., Smolders, J.W.T. & Klinke, R. (1987) *Hear. Res.* **29**, 63–92.
16. Sellick, P.M., Patuzzi, R. & Johnstone, B.M. (1982) *J. Acoust. Soc. Am.* **72**, 131–141.
17. Köppl, C. & Manley, G.A. (1992) in *The Evolutionary Biology of Hearing*, eds. Fay, R.R., Popper, A.N. & Webster, D.B. (Springer, New York), pp. 489–509.
18. Manley, G.A. (2000) in *Comparative Hearing: Birds and Reptiles, Springer Handbook of Auditory Research*, eds. Dooling, R., Popper, A.N. & Fay, R.R. (Springer, New York), in press.
19. Manley, G.A., Köppl, C. & Sneary, M. (1999) *Hear. Res.* **131**, 107–116.
20. Manley, G.A., Köppl, C. & Yates, G.K. (1989) in *Mechanics of Hearing*, eds. Wilson, J.P. & Kemp, D. (Plenum, New York), pp. 143–150.
21. Peake, W.T. & Ling, A. (1980) *J. Acoust. Soc. Am.* **67**, 1736–1745.
22. Holton, T. & Hudspeth, A.J. (1983) *Science* **222**, 508–510.
23. Authier, S. & Manley, G.A. (1995) *Hear. Res.* **82**, 1–13.
24. Manley, G.A., Gleich, O., Kaiser, A. & Brix, J. (1989) *J. Comp. Physiol. A* **164**, 289–296.
25. Gleich, O. & Manley, G.A. (2000) in *Comparative Hearing: Birds and Reptiles, Springer Handbook of Auditory Research*, eds. Dooling, R., Popper, A.N. & Fay, R.R. (Springer, New York), in press.
26. Manley, G.A. (1971) *Nature (London)* **230**, 506–509.
27. Manley, G.A. (1973) *Evolution* **26**, 608–621.
28. Fischer, F.P. (1994) *Scanning Microsc.* **8**, 351–364.
29. Manley, G.A. (1995) in *Advances in Hearing Research*, eds. Manley, G.A., Klump, G.M., Köppl, C., Fastl, H. & Oeckinghaus, H. (World Scientific, Singapore), pp. 219–229.
30. Manley, G.A. (1986) in *Auditory Frequency Selectivity*, eds. Moore, B. & Patterson, R. (Plenum, London), pp. 63–72.
31. Wu, Y.-C., Art, J.J., Goodman, M.B. & Fettiplace, R. (1995) *Prog. Biophys. Mol. Biol.* **63**, 131–158.
32. Köppl, C. (1997) *J. Neurophysiol.* **77**, 364–377.
33. Köppl, C., Manley, G.A. & Konishi, M. (2000) *Curr. Opin. Neurobiol.* in press.
34. Köppl, C., Gleich, O. & Manley, G.A. (1993) *J. Comp. Physiol.* **171**, 695–704.
35. Vater, M. (2000) in *Auditory Worlds: Sensory Analysis and Perception in Animals and Man*, eds. Manley, G.A., Fastl, H., Kössl, M., Oeckinghaus, H. & Klump, G.M. (Wiley, Weinheim), pp. 61–69.
36. Köppl, C. & Manley, G.A. (1990) *J. Comp. Physiol. A* **167**, 101–112.
37. Yates, G.K., Winter, I.M. & Robertson, D. (1990) *Hear. Res.* **45**, 203–220.
38. Liberman, C. & Oliver, M.E. (1984) *J. Comp. Neurol.* **223**, 163–176.
39. Gleich, O. (1989) *Hear. Res.* **37**, 255–268.
40. Yates, G.K., Manley, G.A. & Köppl, C. (2000) *J. Acoust. Soc. Am.* **107**, 2143–2154.
41. Köppl, C. & Yates, G.K. (1999) *J. Neurosci.* **19**, 9674–9686.
42. Davis, H. (1983) *Hear. Res.* **9**, 79–90.
43. Hudspeth, A.J. (1997) *Curr. Opin. Neurobiol.* **7**, 480–486.

44. Manley, G.A. (2000) in *Recent Developments in Auditory Mechanics*, eds. Wada, H., Takasaka, T., Ohyama, K., Ikeda, K. & Koike, T. (World Scientific, Singapore), in press.
45. Köppl, C. (1995) in *Advances in Hearing Research*, eds. Manley, G.A., Klump, G.M., Köppl, C., Fastl, H. & Oeckinghaus, H. (World Scientific, Singapore), pp. 207–216.
46. Benser, M.E., Marquis, R.E. & Hudspeth, A.J. (1996) *J. Neurosci.* **16**, 5629–5643.
47. Denk, W. & Webb, W.W. (1992) *Hear. Res.* **60**, 89–102.
48. Martin, P. & Hudspeth, A.J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 14206–14301.
49. Rüsch, A. & Thurm, U. (1990) *Hear. Res.* **48**, 247–264.
50. Crawford, A.C. & Fettiplace, R. (1985) *J. Physiol.* **364**, 359–379.
51. Steele, C.R. (1996) in *Diversity in Auditory Mechanics*, eds. Lewis, E., Long, G., Lyon, R.F., Narins, P., Steele, C.R. & Hecht-Poinar, E. (World Scientific, Singapore), pp. 455–461.
52. Holley, M.C. (1996) in *The Cochlea*, eds. Dallos, P., Popper, A.N. & Fay, R.R. (Springer, New York), pp. 386–434.
53. Yates, G.K. & Kirk, D.L. (1998) *J. Neurosci.* **18**, 1996–2003.

MECHANICAL BASES OF FREQUENCY TUNING AND NEURAL EXCITATION AT THE BASE OF THE COCHLEA: COMPARISON OF BASILAR-MEMBRANE VIBRATIONS AND AUDITORY-NERVE-FIBER RESPONSES IN CHINCHILLA

Mario A. Ruggero^{*†‡}, S. Shyamla Narayan^{*}, Andrei N. Temchin^{*}, and Alberto Recio[§]

^{*}The Hugh Knowles Center, Department of Communication Sciences and Disorders, and [†]Institute for Neuroscience, Northwestern University, Evanston, IL 60208; and [§]Department of Physiology, University of Wisconsin, Madison, WI 53706

We review the mechanical origin of auditory-nerve excitation, focusing on comparisons of the magnitudes and phases of basilar-membrane (BM) vibrations and auditory-nerve fiber responses to tones at a basal site of the chinchilla cochlea with characteristic frequency ≈ 9 kHz located 3.5 mm from the oval window. At this location, characteristic frequency thresholds of fibers with high spontaneous activity correspond to magnitudes of BM displacement or velocity in the order of 1 nm or 50 $\mu\text{m/s}$. Over a wide range of stimulus frequencies, neural thresholds are not determined solely by BM displacement but rather by a function of both displacement and velocity. Near-threshold, auditory-nerve responses to low-frequency tones are synchronous with peak BM velocity toward scala tympani but at 80–90 dB sound pressure level (in decibels relative to 20 microPascals) and at 100–110 dB sound pressure level responses undergo two large phase shifts approaching 180°. These drastic phase changes have no counterparts in BM vibrations. Thus, although at threshold levels the encoding of BM vibrations into spike trains appears to involve only relatively minor signal transformations, the polarity of auditory-nerve responses does not conform with traditional views of how BM vibrations are transmitted to the inner hair cells. The response polarity at threshold levels, as well as the intensity-dependent phase changes, apparently reflect micromechanical interactions between the organ of Corti, the tectorial membrane and the subreticular fluid, and/or electrical and synaptic processes at the inner hair cells.

In the mammalian cochlea, the bulk of auditory information is transmitted to the brain via the inner hair cells, which provide the sole synaptic inputs to 90–95% of the afferent fibers of the auditory nerve (1). Auditory-nerve excitation is triggered by the depolarization of inner hair cells upon deflection of their “hair” bundles toward the taller “hairs” or stereocilia (2, 3). Presumably, the forces that deflect the stereocilia are derived from the vibrations of the basilar membrane (BM) but it is not known how these vibrations are transmitted to the inner hair cells (4). Although it is clear that the BM and auditory-nerve fibers are similarly tuned at frequencies close to the characteristic frequency (CF) (5, 6), there is no consensus on whether neural thresholds correspond to a constant magnitude of BM displacement, velocity, or some function of these variables. Neither is it known with certainty what phases of BM vibrations trigger auditory-nerve excitation.

Our ignorance on the relationship between the magnitudes and phases of cochlear vibrations and auditory-nerve responses stems in part from the dearth of measurements of BM vibration in relatively healthy cochleae, which until recently were available from a single site in the cochlea of each species. Even when adequate mechanical data are available, comparisons with the responses of auditory-nerve fibers have been perfunctory, typically involving a frequency-threshold tuning curve from a single auditory-nerve fiber in one subject and BM data for another individual of the same species. Given the variability of both neural (e.g., ref. 7) and mechanical responses (e.g., refs. 5 and 6) as well as the difficulty of fully controlling the experimental conditions, such comparisons are bound to be imprecise and inconclusive.

Here we review the cochlear mechanical bases of auditory-nerve excitation as revealed by comparisons of the magnitudes and phases of BM and auditory-nerve-fiber responses to tones recorded from a basal site of the chinchilla cochlea with CF ≈ 9 kHz located about 3.5 mm from the oval window. To date, this site is the only location for which systematic and extensive sets of BM and auditory-nerve-fiber responses are available in any species. Using results from the literature as well as previously unpublished data we address three issues, namely frequency tuning at threshold levels, the timing of auditory-nerve excitation at threshold, and the intensity dependence of auditory-nerve response phases. These issues are investigated by using two complementary approaches. One approach consists of comparing averages of BM and auditory-nerve fiber data (magnitudes or phases) obtained from separate sets of normal subjects. The second approach consists of carrying out the comparisons by using BM and neural data collected consecutively in the same individual cochleae, under identical conditions.

Methods

The data of Figs. 1, 3A, 5, and 6 have not been previously published. The new BM and auditory-nerve data were obtained by using methods described, respectively, in refs. 8 and 9. For the purpose of relating BM and neural response phases, the latter have been corrected for an estimated 1-ms (frequency independent) delay introduced by synaptic processes and neural conduction time (for justification see refs. 9 and 10). Therefore, the corrected neural phases of Figs. 4–6 indicate, to a good first approximation, the phases of depolarization of inner hair cells.

Frequency Tuning of BM and Auditory-Nerve Responses to Tones

Fig. 1 summarizes the responses to tones of a BM site located 3.5 mm away from the oval window in a nearly normal chinchilla

[‡]To whom reprint requests should be addressed at: Communication Sciences and Disorders, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208–3550. E-mail: mruggero@northwestern.edu.

This paper was presented at the National Academy of Sciences Colloquium “Auditory Neuroscience: Development, Transduction and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: BM, basilar membrane; CF, characteristic frequency; SPL, sound pressure level.

cochlea. Vibration magnitudes (Fig. 1 A and B) and phases (Fig. 1 C and D) are plotted as a function of frequency (abscissa) and stimulus level (ordinate). The magnitudes grow at compressive rates at frequencies near CF (9.5 kHz) but linearly at other frequencies, with the result that sensitivity at CF decreases and frequency tuning becomes broader as a function of increasing stimulus level (Fig. 1 A and B). Response phases increasingly lag as stimulus frequency increases, with a relatively shallow slope at low frequencies and steep slopes near CF (Fig. 1 C), a frequency region within which phases shift systematically with stimulus level (Fig. 1 D). At frequencies well apart from CF, response phases do not vary with stimulus levels.

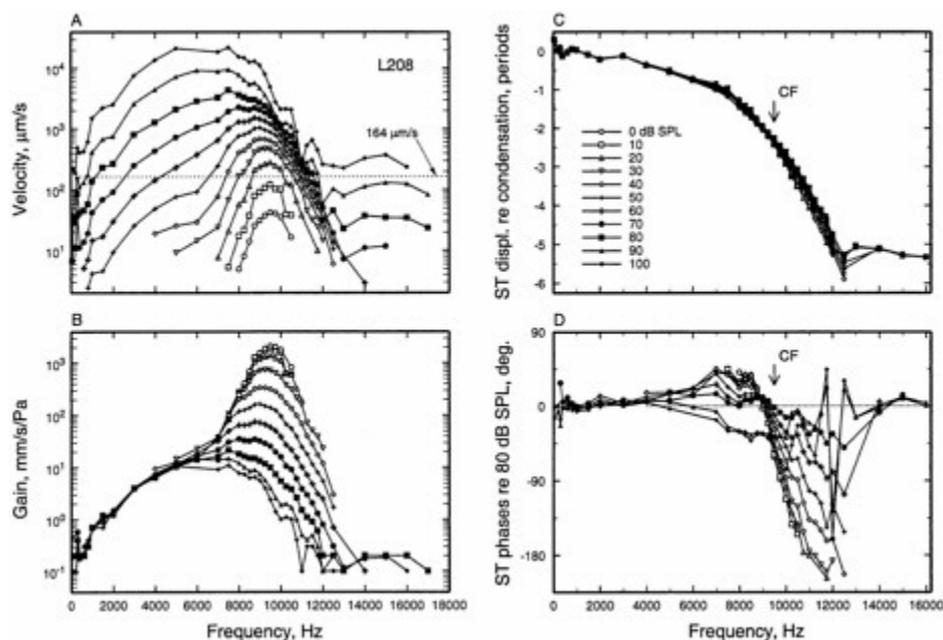


Fig. 1. The magnitudes and phases of BM responses to tones at the 3.5-mm site of the chinchilla cochlea. (A) Velocity magnitude as a function of stimulus frequency (abscissa) and level (parameter). The dotted line indicates the velocity corresponding to the CF threshold of the auditory-nerve fiber in Fig. 3B. (B) As in A, but normalized to stimulus pressure. (C) Response phases of peak displacement toward scala tympani, relative to peak condensation at the external ear canal, as a function of stimulus frequency and level. (D) As in C, but normalized to the phases of responses to 80-dB stimuli. Data from cochlea L208, recorded by using a laser vibrometer (59).

To compare frequency tuning at various stages of auditory-signal transduction it is convenient and customary to construct equal-response curves (e.g., Fig. 2) expressing the sound pressure level (SPL) required to produce a constant response magnitude. In the case of auditory-nerve fibers, frequency tuning usually is specified in terms of threshold. Such frequency-threshold curves may be compared with equal-velocity or equal-displacement curves computed from BM responses (e.g., Fig. 2) by determining the SPLs required to elicit a given magnitude of velocity (e.g., dotted line in Fig. 1A) or displacement. The first explicit comparison between neural thresholds and essentially normal BM vibrations was performed for a site of the guinea pig cochlea with CF \approx 18 kHz (5). The thresholds for a single auditory-nerve fiber recorded in one cochlea were thought to correspond to a fixed magnitude of BM velocity, about 40 μ m/s (5). However, on the basis of the same data it was later argued that neural thresholds correspond to a fixed BM displacement (11). A comparison for responses from the hook region of the cat cochlea led to a similar conclusion, with neural threshold corresponding to BM vibrations amounting to about 1 nm (12).

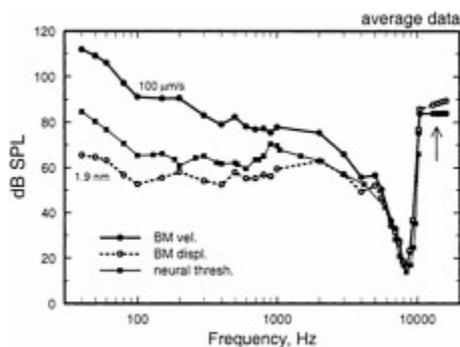


Fig. 2. Frequency tuning of BM vibrations and auditory-nerve fibers at the 3.5-mm site of the chinchilla cochlea. An average frequency-threshold curve computed from responses of many auditory-nerve fibers is compared with the average tuning of BM responses in several cochleae, expressed as the stimulus levels at which BM vibration attains a displacement of 1.9 nm or a velocity of 100 μ m/s. BM data were measured by using the Mössbauer technique. Redrawn from the data of figure 22 of Ruggero et al. (17).

In the chinchilla, average tuning curves for auditory-nerve fibers were compared with BM data recorded at sites with CFs of 9 and 10 kHz in two individual cochleae (13). Near CF there was an excellent match between neural and mechanical tuning, with neural thresholds corresponding to 39 and 73 $\mu\text{m/s}$ (or, equivalently, 0.62 and 1.3 nm). However, because the comparisons were carried out over a limited range of frequencies (>2 kHz), no firm conclusion was reached regarding the precise relationship between mechanical and neural tuning.

Mechanical tuning also has been compared with tuning curves for inner hair cells. Sellick *et al.* (14) compared BM isovelocity and isodisplacement curves from guinea pig cochleae with one inner hair cell tuning curve (based on a constant DC receptor potential). Perusal of those data does not suggest any obvious conclusion on whether isovelocity or isodisplacement provides a better match to the receptor potentials. However, another study in guinea pig (15) found that tip-to-tail ratios for the receptor potential of one inner hair cell were substantially higher than those for displacement magnitude of BM vibrations.

Comparison of Averaged Mechanical and Neural Responses at the 3.5-mm Site of the Chinchilla Cochlea. The problems inherent in comparing data from different individual subjects are ameliorated by making comparison using averages of grouped data. To date, such comparisons have been carried out only for the 3.5-mm site of the chinchilla cochlea (Fig. 2). An average frequency-threshold curve, computed from responses of many auditory-nerve fibers recorded in one group of chinchillas, was compared with tuning curves computed for velocity and displacement of (nearly normal) BM responses recorded with the Mössbauer technique in another group of chinchillas. The comparison indicated that, at CF (≈ 8.5 kHz), neural threshold corresponded to a displacement of about 2 nm or, equivalently, 100 $\mu\text{m/s}$. However, when the comparison was carried out over a range of stimulus frequencies encompassing more than two decades (over which the magnitudes of velocity and displacement diverge widely, at a rate of 6 dB per octave or 20 dB per decade), the frequency tuning of auditory-nerve fibers did not match a constant BM displacement or velocity, but rather corresponded to intermediate values (6, 16, 17).

An updated comparison of averaged neural and BM tuning for the same site of the chinchilla cochlea now can be carried out by using new BM measurements obtained with a laser vibrometer, as well as a larger database of frequency-threshold curves for auditory-nerve fibers. We want to ascertain that the conclusion drawn from Fig. 2 was unsullied by inclusion of BM responses that might have been abnormally insensitive at near-CF frequencies because of cochlear damage. Therefore, although it is desirable to compare mechanical and neural responses over the widest possible range of stimulus frequencies, we limit the present comparison to frequencies lower than 5 kHz, at which BM responses are linear and resistant to cochlear insults (5, 6, 13). The lower limits of the frequency range of measurements are dictated by the insensitivity of both neural and mechanical responses to low-frequency stimuli and, in the case of laser velocimetry recordings, by the presence of an artifact (because of motion of the perilymph meniscus overlying the BM measurements site) whose effects increase with decreasing frequency (12, 13, 18).

Average neural thresholds and BM vibration magnitudes as a function of frequency are compared in Fig. 3A. Average rate thresholds were computed for auditory-nerve fibers with spontaneous activity >18 spikes/s and CFs in the range of 8 to 12 kHz. Fibers with high spontaneous activity were selected because, in contrast with fibers with medium or low spontaneous activity, their thresholds are uniform in any individual ear at any given CF region (7). BM vibration magnitudes are presented in Fig. 3A as curves of constant velocity or constant displacement, which were arbitrarily equated at 5 kHz. At this frequency, neural threshold corresponds to 1.5 nm or 47 $\mu\text{m/s}$. Fig. 3A makes it clear that, in the range of 200 to 5,000 Hz, neural thresholds do not correspond to a constant BM displacement. Rather, neural thresholds closely approximate displacement magnitudes that have undergone high-pass filtering at a rate of $+3.83$ dB/octave.

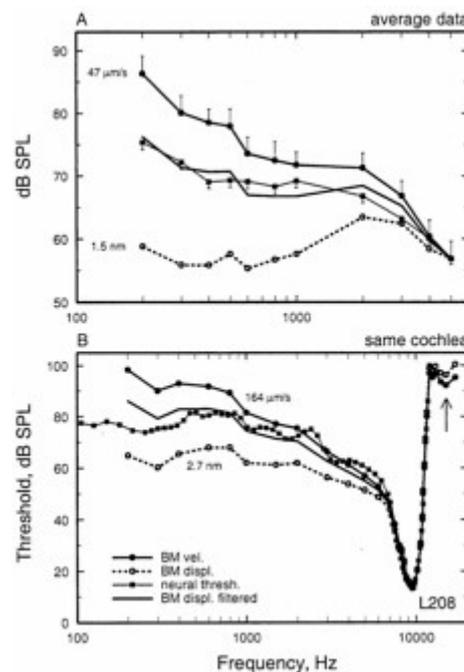


Fig. 3. Frequency tuning of BM vibrations and auditory-nerve fibers at the 3.5-mm site of the chinchilla cochlea. (A) An average frequency-threshold curve for auditory-nerve fibers is compared with the average tuning of BM responses in several cochleae, expressed as three curves indicating the stimulus levels at which BM vibration attains a displacement of 1.5 nm, a velocity of 47 $\mu\text{m/s}$, and a constant displacement high-pass filtered at a rate of 3.83 dB/octave. The three curves are equated to neural threshold at 5 kHz. BM data were measured in 18 cochleae by using laser velocimetry. The threshold averages of auditory-nerve fibers were based on data for 183 fibers with CF= 8–12 kHz recorded in 77 chinchillas. The data consisted of frequency-threshold tuning curves, measured with an automated adaptive algorithm (7, 60), and 0.5-s samples of responses to low-frequency tones (≤ 1 kHz) (9). Vertical bars indicate the SEM. (B) Frequency tuning of responses to tones of a BM site and an auditory-nerve fiber with similar characteristic frequency recorded in the same cochlea. The neural frequency-threshold tuning curve is compared with BM tuning curves indicating constant displacement (2.7 nm), constant velocity (164 $\mu\text{m/s}$), and displacement high-pass filtered at a rate of 3.81 dB/octave. The auditory-nerve fiber had spontaneous activity of 11.2 spikes/s. (B) Redrawn from figure 1B of Narayan *et al.* (8).

Comparison of Mechanical and Neural Responses Recorded in the Same Chinchilla Cochleae. In two exceptional cases, we succeeded in recording from auditory-nerve fibers and BM sites with corresponding CFs in the same chinchilla cochleae (8). The BM data of Fig. 1 were obtained from one of these cochleae, in which four auditory-nerve fibers were encountered with CFs (7.8–9.5

kHz) comparable to the CF of the BM recording site (9.5 kHz). Comparison of responses is especially straightforward in the case of the 9.5-kHz fiber (Fig. 3B), which presumably innervated an inner hair cell immediately adjacent to the BM recording site. At the fiber's CF threshold (13 dB SPL), BM vibrations had a peak velocity of 164 $\mu\text{m/s}$ (dotted line in Fig. 1A) or, equivalently, a peak displacement of 2.7 nm. These values were used to specify mechanical isodisplacement and isovelocity tuning curves. At frequencies between CF and 1 kHz, there was a good match between neural thresholds and a constant BM velocity. However, taking into account the entire frequency range of measurements, neural thresholds were better fit by mechanical displacements subjected to high-pass filtering at a rate of 3.81 dB/octave. The tuning curves for the three other fibers with similar CF were well fit by BM displacement high-pass filtered at rates of 3.9–4.1 dB/octave. A similar comparison was carried out in another cochlea, in which the BM recording site had a CF of 9 kHz and four fibers were found with comparable CFs. The tuning curves of these fibers also were well matched by high-pass filtered BM displacement (at rates of 2.7–6 dB/octave).

Thus, at near-threshold stimulus levels, the frequency tuning of auditory-nerve fibers in both cochleae closely resembled that of BM displacement modified by high-pass filtering. However, neural tuning curves lacked the high-frequency plateaus (arrows in Figs. 2 and 3B) usually demonstrable in BM responses (6, 12, 17, 19, 20), suggesting that such vibrations are not transmitted to the stereocilia of inner hair cells (21).

Timing of Auditory-Nerve Threshold Responses to Low-Frequency Tones in Relation to BM Motion

The traditional view of auditory-nerve excitation was that spikes are generated when the BM is displaced toward scala vestibuli (22). This view was based on the functional polarity of the hair cell stereociliar bundles (which indicates that depolarization should occur when stereocilia are deflected away from the modiolus) and the likelihood that shear must be generated between the reticular lamina and the tectorial membrane sites contacted by the tips of the stereocilia (23, 24). An updated version of this view, partly based on the likelihood that inner hair cell stereocilia are not firmly connected to the tectorial membrane (25, 26), is that inner hair cells are depolarized (and auditory-nerve fibers excited) when the BM is moving at maximum speed from scala tympani toward scala vestibuli (27, 28). At the apex of the cochlea, both intracellular recordings from guinea pig inner hair cells (29) as well as responses of chinchilla auditory-nerve fibers (9, 10, 30) support the notion that depolarization of the inner hair cells occurs when the cochlear partition is in motion toward scala vestibuli.

A rather different situation holds for basal sites of the cochlea, where recordings from inner hair cells and auditory-nerve fibers are in conflict. As at the apex, depolarization responses of basal inner hair cells to low-frequency stimuli are in phase with peak BM velocity toward scala vestibuli (31–33). In contrast, responses of auditory-nerve fibers to low-frequency, nearthreshold stimulation appear to be synchronous with BM displacement or motion toward scala tympani. Traditional views of inner hair cell stimulation were first challenged by studies of the responses of gerbil and guinea pig auditory-nerve fibers to very low-frequency stimuli (<100 Hz), designed to produce displacement steps or trapezoids at the BM (34–36). These studies, which deduced BM displacement from cochlear microphonics recorded at the round window, concurred that the timing of auditory-nerve fiber excitation differs significantly between basal and apical cochlear locations and suggested that fibers innervating the cochlear base are stimulated when the BM is displaced or in motion toward scala tympani. Later investigations, although often disagreeing in many details and also relying on cochlear microphonics to deduce BM motion, generally supported these conclusions, which were found to also apply to tonal stimuli with much higher frequencies (30, 37, 38).

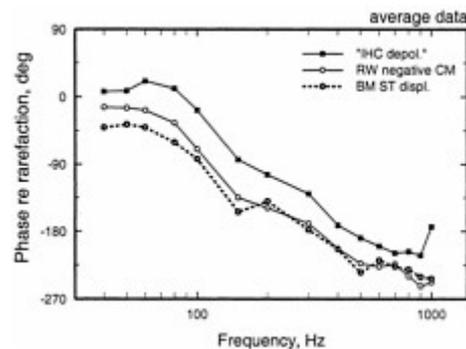


Fig. 4. Average response phases of auditory-nerve fibers, cochlear microphonics (recorded at the round window), and BM displacement at the base of the chinchilla cochlea. Redrawn from the data of figure 5 of Ruggero *et al.* (16).

Comparison of Averaged Phases for Mechanical and Neural Responses Recorded at the Base of the Cochlea. The first recordings of BM vibration in the chinchilla cochlea, obtained with the Mössbauer technique, yielded response phases for low-frequency stimuli that were consistent with the phases of cochlear microphonics recorded at the round window (Fig. 4) (30, 39). Fig. 4 compares the phases of BM vibrations and auditory-nerve fiber excitation in response to low frequency (40–1,000 Hz) tones presented at near-threshold levels (16). [Comparisons are restricted to responses to low-frequency stimuli because phase locking in the auditory nerve diminishes rapidly above 1 kHz (40, 41) and also because instrumental and neural jitter greatly disturb the determination of phase at higher frequencies.] According to Fig. 4, neural excitation is triggered at an instant intermediate between the times of peak BM displacement and velocity toward scala tympani (i.e., leading BM displacement toward scala tympani; Fig. 4) (6, 10, 17). This response polarity was surprising because it conflicted with the phases of inner hair cell depolarization responses to low-frequency tones at the base of the guinea pig cochlea (31, 32). The paradoxical difference between the polarity of responses of inner hair cells and auditory-nerve fibers is not caused by species differences, because it also holds for inner hair cell and auditory-nerve data recorded at the base of guinea pig cochlea (38).

Fig. 5A shows average phases for the BM responses (of 8–18 cochleae, depending on frequency, newly recorded by using laser velocimetry) whose magnitudes are shown in Fig. 3A, together with average phases of near-threshold responses of auditory-nerve fibers with CFs of 8–12 kHz taken from a database of Ruggero *et al.* (9). Throughout the 200- to 1,000-Hz range of stimulus frequencies, peak inner hair cell depolarization (deduced from the timing of neural excitation) was nearly synchronous with peak BM velocity toward scala tympani, confirming the response polarity of Fig. 4 but indicating a somewhat larger phase lead relative to BM displacement.

Comparison of BM and Auditory-Nerve Fiber Responses Recorded at the Base of the Same Cochleae. Fig. 5B compares the phases of neural and mechanical responses to tones in one of two cochleae from which responses were recorded from both the BM and auditory nerve fibers (8). Among the latter, three had CFs

(7.1–8.1 kHz) similar to the CF (9 kHz) of the BM recording site. At stimulus frequencies <1 kHz, the neural response phases matched closely those of peak BM velocity toward scala tympani, consistent with the comparison of averaged data of Fig. 5A. Results in the other cochlea in which both neural and BM data were recorded (represented in Figs. 1 and 3B) were similar to those of Fig. 5B.

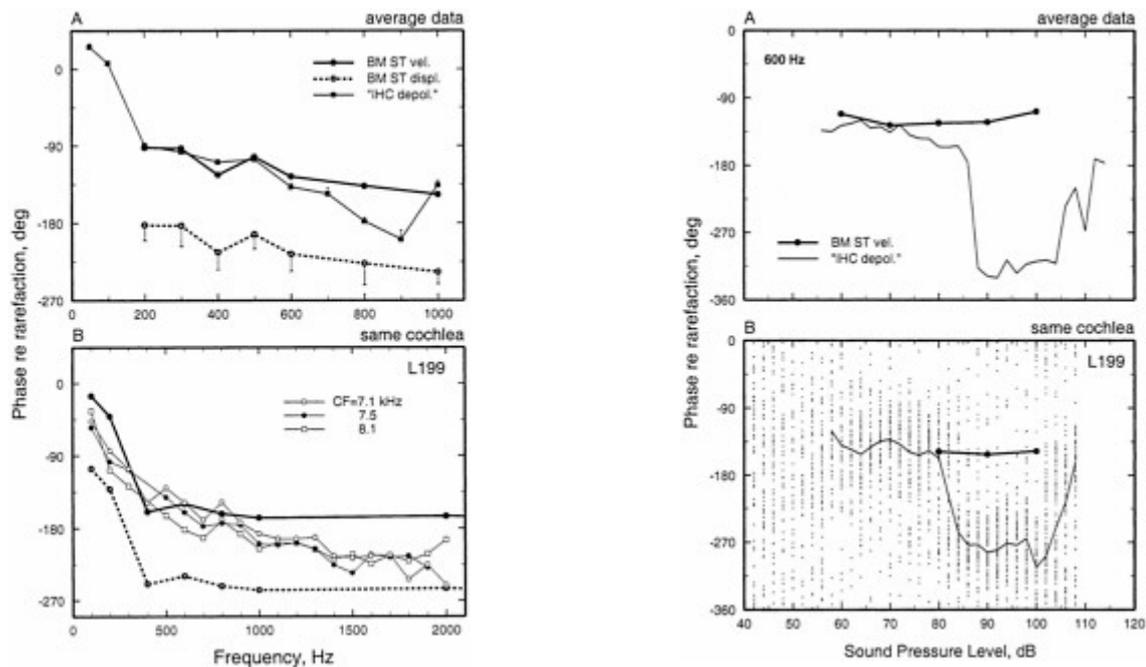


Fig. 5. The phases of BM and auditory-nerve-fiber responses to low-frequency tones. (A) Average response phases of auditory-nerve fibers and BM vibrations at the base of the chinchilla cochlea. Neural data are averages from 13–52 fibers (depending on frequency) with CF 8–12 kHz. BM responses were measured in 8–18 cochleae (depending on frequency). Vertical bars indicate SEM. (B) The phases of responses to tones of a BM site and auditory-nerve fibers with similar CF recorded in the same cochlea. The auditory-nerve phases have been corrected for 1 ms (the sum of neural and synaptic delays) so that they indicate the presumed phases of peak depolarization of the inner hair cells (9). The same BM data are represented in two curves, depicting the phases of peak velocity toward scala tympani and of peak displacement toward scala tympani. Spontaneous rates were 105, 62, and 73 spikes/s, respectively. Thresholds at CF were 11, 4, and 1 dB SPL, respectively. Stimuli for the auditory-nerve fibers were 10 repetitions of 75-dB, 100-ms tone pips presented every 300 ms. The data points for BM phases represent averages for responses to 50–100 dB SPL, depending on frequency. Averaging is justified because, as shown in Fig. 6B, BM response phases did not vary significantly as a function of stimulus intensity.

Fig. 6. The phases of BM and auditory-nerve-fiber responses to low-frequency tones as a function of stimulus level. (A) Average phases of responses to 600-Hz tones recorded at the BM of eight cochleae and from 27 auditory-nerve fibers with CF 8–12 kHz. (B) The responses to 600-Hz tones of a BM site and an auditory-nerve fiber (CF=7.1 kHz; spontaneous rate=105 spikes/s) with similar CF recorded in the same cochlea. The phases of peak BM velocity toward scala tympani are indicated by \sim connected by a thick line. The neural responses are depicted as a scatter diagram of phase vs. level, with each dot representing one spike. Stimuli were five repetitions of 100-ms tones, presented every 300 ms at random levels, with 2-dB steps. The mean phases of the neural responses are indicated by a thin line.

Intensity Dependence of Response Phases of Auditory-Nerve Fibers

Many years ago, Nelson Kiang and colleagues described striking intensity-dependent irregularities in the responses of cat auditory-nerve fibers (40, 42). These consisted of abrupt phase shifts, a bimodal distribution of excitation within each period of the response to a low frequency tone (peak splitting), and a notch in rate-intensity functions. Similar phenomena have been subsequently studied in cat (43–46), guinea pig (38), and chinchilla (9, 30). In chinchilla the basic phenomenon, a 90–180° phase shift, occurs almost universally, at 85–90 dB SPL, in responses of auditory-nerve fibers to low-frequency tones. An example of the phase shift for responses to 600-Hz tones is shown in Fig. 6B. At the exact intensity of the phase shift, period histograms also may exhibit peak splitting and, less often, response rates may dip sharply. Interpretation of these phenomena usually has been based on postulating the existence of two modes of excitation, which grow at different rates, have different phases, and can cause mutual cancellation (44, 45). In the case of high-CF auditory-nerve fibers, the possibility that similar phase shifts might take place in BM vibrations has been considered but usually rejected because of the linear nature of basal BM responses to stimuli with frequency well below CF (see Fig. 1). Nevertheless, until now we did not consider the matter closed because the existence of such phase shifts in BM responses was seldom addressed explicitly and was not well tested. In particular, BM recordings using the Mössbauer technique rarely permitted obtaining reliable responses spanning the intensity range over which neural phase shifts occur, because of its relative insensi

tivity combined with the insensitivity of BM responses to low-frequency tones at basal (high-CF) sites. We recently have revisited this subject by using a more sensitive technology (laser velocimetry).

Comparison of Averaged Phases for Mechanical and Neural Responses Recorded at the Base of the Cochlea. Fig. 6A displays the average phases of responses to 600-Hz tones of 27 chinchilla auditory-nerve fibers with CFs of 8–12 kHz, plotted as a function of stimulus intensity. Response phases are constant over a wide range of intensities, from threshold to 86 dB SPL, but undergo a reversal as the intensity is stepped from 86 to 88 dB SPL. Phases stay constant in the range of 88 to 105 dB but undergo another reversal at 105–110 dB. Similar shifts were found for stimulus frequencies 100–1,000 Hz (9). For comparison, Fig. 6A also shows average phases for responses of a basal BM site with CF of 9–10 kHz to 600-Hz tones presented at 60, 70, 80, 90, and 100 dB SPL. Although absolute phases varied substantially from cochlea to cochlea (not shown), in any single cochlea, for a given stimulus frequency, the variation of BM phase with intensity was insignificant (average standard deviation: 7°). The slight variations with intensity of the average phases of Fig. 6A reflects almost entirely a combination of the variation of response phase across cochleae and differences in the sample size across stimulus levels. Comparison for other stimulus frequencies in the range of 100 to 1,000 Hz yielded similar results in that the neural responses underwent phase reversals but BM responses did not.

The Phases of BM and Auditory-Nerve Fiber Responses Recorded at the Base of the Same Cochleae. A stringent test for the existence of BM counterparts of the intensity phase shifts seen in neural responses was carried out by recording BM and auditory-nerve-fiber responses to the same stimuli in the same cochleae (Fig. 6B). Fig. 6B displays the responses to 600-Hz tones of a fiber (CF: 7.1 kHz) recorded in one of these cochleae, as a function of stimulus intensity. Each dot represents an action potential. At intensities between threshold and 80 dB, most discharges were synchronous with peak BM velocity toward scala tympani. At 80–84 dB, the timing of discharge shifted abruptly, by about 140°. A second abrupt shift took place at about 100 dB, with phases returning to their near-threshold values. In contrast with these phase shifts, the phases of BM vibration did not change significantly as a function of stimulus intensity. Similar findings were obtained for stimulus frequencies 200–800 Hz in other fibers in the same and in another cochlea (8).

Summary Discussion

Systematic comparisons of the responses of auditory-nerve fibers and BM vibrations have been carried out only for the 3.5-mm site of the chinchilla cochlea. Thus, it remains to be seen whether the present findings apply to other sites of the cochlea of the chinchilla as well as other species.

Frequency Tuning. The similar frequency tuning exhibited by BM vibrations and auditory-nerve fiber thresholds at the 3.5-mm site of the chinchilla cochlea refutes lingering arguments in defense of the existence in mammalian cochleae of a “second filter” (e.g., refs. 47 and 48), analogous to the process that sharpens the frequency tuning of hair cells in the basilar papilla of turtles (49, 50). However, the comparisons of BM and neural response magnitudes and phases suggest that frequency-threshold neural tuning curves are not simply determined by BM displacement, but rather by a function of both BM displacement and velocity. A role of BM velocity in neural excitation is reasonable if the stereocilia of inner hair cells are not attached to the tectorial membrane (25, 26), because in that case deflection of the stereocilia would be accomplished by their motion relative to the surrounding fluid (27, 28). Such a velocity dependence of receptor potentials has been demonstrated in basal inner hair cells, with an upper-frequency limit of velocity sensitivity perhaps as high as 1,600 Hz (see figure 7 of ref. 33). Nevertheless, it is prudent to think of the terms displacement and velocity more as shorthand expressions rather than as literal indicators of how inner hair cell stereocilia are stimulated, because it is likely that more central stages of cochlear signal processing (i.e., transduction currents, receptor potentials, calcium currents, transmitter release, and spike generation) act as (or mimic the effects of) frequency filters.

Polarity of Auditory-Nerve Fiber Excitation. The polarity of auditory-nerve fiber excitation relative to BM vibrations at the base of the cochlea remains the most puzzling of our findings. First, the polarity of auditory-nerve fiber excitation at threshold levels seems to be incompatible with recordings of inner hair cell receptor potentials. This discrepancy may reflect dissimilar stimulation conditions (e.g., stimulus intensity) or may indicate that the recordings from inner hair cells are flawed (e.g., because the microelectrode interferes with cochlear micromechanics; ref. 51). Second, the polarity of excitation at threshold is opposite that predicted by the standard assumption that the stereocilia of hair cells (including inner hair cells) are deflected toward the taller stereocilia (the depolarizing direction) by rotation of the reticular lamina toward scala vestibuli (23, 24). Several schemes have been put forth to explain the response phases of basal auditory-nerve fibers, including their dependence on stimulus level, on the basis of micromechanics, electrical filtering of receptor potentials by the basolateral-membrane of inner hair cells, and a possible influence of extracellular microphonics (9, 32, 38, 52, 53). However, it seems fair to state that none have succeeded in accounting for the full complexity of the findings [see reviews by Ruggero *et al.* (9) and Cheatham and Dallos (53)].

Intensity Dependence of Auditory-Nerve Fiber Response Phases. The drastically different behaviors of BM and neural response phases as a function of stimulus level rule out the possibility that the intensity-dependent phase shifts of auditory-nerve fibers reflect multiple modes in BM vibrations (54–57) but suggest that multiple vibration modes exist in the micromechanics of the organ of Corti/tectorial membrane complex (9, 44, 45, 58).

We were supported by Grant DC-00419 from the National Institute on Deafness and Other Communication Disorders.

1. Spoendlin, H. (1988) in *Physiology of the Ear*, eds. Jahn, A.F. & Santos-Sacchi, J. (Raven, New York), pp. 201–219.
2. Hudspeth, A.J. & Corey, D.P. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2407–2411.
3. Russell, I.J., Richardson, G.P. & Cody, A.R. (1986) *Nature (London)* **321**, 517–519.
4. Ruggero, M.A. (1995) in *Active Hearing*, eds. Flock, A., Ottoson, D. & Ulfendahl, M. (Pergamon, Oxford), pp. 321–336.
5. Sellick, P.M., Patuzzi, R. & Johnstone, B.M. (1982) *J. Acoust. Soc. Am.* **72**, 131–141.
6. Robles, L., Ruggero, M.A. & Rich, N.C. (1986) *J. Acoust. Soc. Am.* **80**, 1364–1374.
7. Liberman, M.C. (1978) *J. Acoust. Soc. Am.* **63**, 442–455.
8. Narayan, S.S., Temchin, A.N., Recio, A. & Ruggero, M.A. (1998) *Science* **282**, 1882–1884.
9. Ruggero, M.A., Rich, N.C., Shivapuja, B.G. & Temchin, A.N. (1996) *Audit. Neurosci.* **2**, 159–185.
10. Ruggero, M.A. & Rich, N.C. (1987) *J. Neurophysiol.* **58**, 379–403.
11. Neely, S.T. & Kim, D.O. (1983) *Hear. Res.* **9**, 123–130.
12. Cooper, N.P. & Rhode, W.S. (1992) *Hear. Res.* **63**, 163–190.
13. Ruggero, M.A., Rich, N.C., Recio, A., Narayan, S.S. & Robles, L. (1997) *J. Acoust. Soc. Am.* **101**, 2151–2163.
14. Sellick, P.M., Patuzzi, R. & Johnstone, B.M. (1983) *Hear. Res.* **10**, 93–100.

15. Russell, I.J., Kössl, M. & Murugasu, E. (1995) in *Advance in Hearing Research*, eds. Manley, G.A., Klump, G.M., Köppl, C., Fastl, H. & Oeckinghaus, H. (World Scientific, Singapore), pp. 136–144.
16. Ruggero, M.A., Robles, L. & Rich, N.C. (1986) *J. Acoust. Soc. Am.* **80**, 1375–1383.
17. Ruggero, M.A., Rich, N.C., Robles, L. & Shivapuja, B.G. (1990) *J. Acoust. Soc. Am.* **87**, 1612–1629.
18. Recio, A., Rich, N.C., Narayan, S.S. & Ruggero, M.A. (1998) *J. Acoust. Soc. Am.* **103**, 1972–1989.
19. Rhode, W.S. (1971) *J. Acoust. Soc. Am.* **49**, 1218–1231.
20. Nuttall, A.L. & Dolan, D.F. (1996) *J. Acoust. Soc. Am.* **99**, 1556–1565.
21. Wilson, J.P. (1992) in *Auditory Physiology and Perception*, eds. Cazals, Y., Demany, L. & Horner, K. (Pergamon, Oxford), pp. 71–84.
22. Goldstein, M.H., Jr. (1968) in *Medical Physiology*, ed. Mouncastle, V.B. (Mosby, St. Louis), pp. 1465–1498.
23. ter Kuile, E. (1900) *Pflügers Arch. Ges. Physiol. Menschen Tiere* **79**, 146–157.
24. Davis, H. (1965) *Cold Spring Harb. Symp. Quant. Biol.* **30**, 181–189.
25. Lim, D. (1980) *J. Acoust. Soc. Am.* **67**, 1686–1695.
26. Lim, D. (1986) *Hear. Res.* **22**, 117–146.
27. Dallos, P., Billone, M.C., Durrant, J.D., Wang, C.-Y. & Raynor, S. (1972) *Science* **177**, 356–358.
28. Billone, M. & Raynor, S. (1973) *J. Acoust. Soc. Am.* **54**, 1143–1156.
29. Santos-Sacchi, J. & Dallos, P. (1983) *Hear. Res.* **9**, 317–326.
30. Ruggero, M.A. & Rich, N.C. (1983) *J. Acoust. Soc. Am.* **73**, 2096–2108.
31. Nuttall, A.L., Brown, M.C., Masta, R.I. & Lawrence, M. (1981) *Brain Res.* **211**, 171–174.
32. Russell, I.J. & Sellick, P.M. (1983) *J. Physiol. (London)* **338**, 179–206.
33. Patuzzi, R. & Yates, G. (1987) *Hear. Res.* **30**, 83–98.
34. Konishi, T. & Nielsen, D.W. (1978) *Jpn. J. Physiol.* **28**, 291–307.
35. Zwislocki, J.J. & Sokolich, W.G. (1973) *Science* **182**, 64–66.
36. Sokolich, W.G., Hamernik, R.P., Zwislocki, J.J. & Schmiedt, R.A. (1976) *J. Acoust. Soc. Am.* **59**, 963–974.
37. Oshima, W. & Strelhoff, D. (1983) *Hear. Res.* **12**, 167–184.
38. Sellick, P.M., Patuzzi, R. & Johnstone, B.M. (1982) *Hear. Res.* **7**, 199–222.
39. Ruggero, M.A., Robles, L. & Rich, N.C. (1986) *J. Acoust. Soc. Am.* **79**, 1491–1498.
40. Johnson, D.H. (1980) *J. Acoust. Soc. Am.* **68**, 1115–1122.
41. Weiss, T. & Rose, C. (1988) *Hear. Res.* **33**, 167–174.
42. Kiang, N.Y.S. & Moxon, E.C. (1972) *Ann. Otol. Rhinol. Laryngol.* **81**, 714–730.
43. Gifford, M.L. & Guinan, J.J., Jr. (1983) *J. Acoust. Soc. Am.* **74**, 115–123.
44. Kiang, N.Y.S. (1984) in *Handbook of Physiology—The Nervous System: Sensory Processes*, ed. Darian-Smith, I. (Am. Physiol. Soc., Bethesda, MD), pp. 639–674.
45. Kiang, N.Y.S. (1990) *Hear. Res.* **49**, 1–16.
46. Liberman, M.C. & Kiang, N.Y.S. (1984) *Hear. Res.* **16**, 75–90.
47. Allen, J.B. & Neely, S.T. (1992) *Physics Today* **45**, 40–47.
48. Allen, J.B. & Fahey, P.F. (1993) *J. Acoust. Soc. Am.* **94**, 809–816.
49. Crawford, A.C. & Fettiplace, R. (1981) *J. Physiol. (London)* **312**, 377–412.
50. Art, J.J. & Fettiplace, R. (1987) *J. Physiol. (London)* **385**, 207–242.
51. Zwislocki, J.J. (1985) *Acta Oto-Laryngol.* **100**, 201–209.
52. Cheatham, M.A. & Dallos, P. (1998) *J. Acoust. Soc. Am.* **104**, 356–369.
53. Cheatham, M.A. & Dallos, P. (1999) *J. Acoust. Soc. Am.* **105**, 799–810.
54. Kolston, P.J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 3676–3681.
55. Hubbard, A. (1993) *Science* **259**, 68–71.
56. Mountain, D.C. & Cody, A.R. (1999) *Hear. Res.* **132**, 1–14.
57. Nilsen, K.E. & Russell, I.J. (1999) *Nat. Neurosci.* **2**, 642–648.
58. Lin, T. & Guinan, J.J., Jr. (2000) *J. Acoust. Soc. Am.* **107**, 2615–2630.
59. Ruggero, M.A. & Rich, N.C. (1991) *Hear. Res.* **51**, 215–230.
60. Wiederhold, M.L. & Kiang, N.Y. (1970) *J. Acoust. Soc. Am.* **48**, 950–965.

THE SPATIAL AND TEMPORAL REPRESENTATION OF A TONE ON THE GUINEA PIG BASILAR MEMBRANE

K.E.Nilsen and I.J.Russell*

School of Biological Sciences, University of Sussex, Falmer Brighton, BN1 9QG, United Kingdom

In the mammalian cochlea, the basilar membrane's (BM) mechanical responses are amplified, and frequency tuning is sharpened through active feedback from the electromotile outer hair cells (OHCs). To be effective, OHC feedback must be delivered to the correct region of the BM and introduced at the appropriate time in each cycle of BM displacement. To investigate when OHCs contribute to cochlear amplification, a laser-diode interferometer was used to measure tone-evoked BM displacements in the basal turn of the guinea pig cochlea. Measurements were made at multiple sites across the width of the BM, which are tuned to the same characteristic frequency (CF). In response to CF tones, the largest displacements occur in the OHC region and phase lead those measured beneath the outer pillar cells and adjacent to the spiral ligament by about 90°. Postmortem, responses beneath the OHCs are reduced by up to 65 dB, and all regions across the width of the BM move in unison. We suggest that OHCs amplify BM responses to CF tones when the BM is moving at maximum velocity. In regions of the BM where OHCs contribute to its motion, the responses are compressive and nonlinear. We measured the distribution of nonlinear compressive vibrations along the length of the BM in response to a single frequency tone and estimated that OHC amplification is restricted to a 1.25- to 1.40-mm length of BM centered on the CF place.

cochlear amplifier | outer hair cell | frequency selectivity | laser-diode interferometry

Von Békésy (1) discovered that sound-induced stapes movement causes a wave of basilar membrane (BM) displacement to travel from the base of the cochlea to the apex because of a pressure difference set up across the cochlear partition. The length of the BM is graded in stiffness, and the cochlear partition can be modeled as a series of weakly coupled sections, each section comprised of a rigid mass connected to the sides of the cochlea by springs that decrease in stiffness toward the apex (2). The BM will resonate when the mechanical impedance caused by the stiffness of the springs cancels that caused by the mass of the section (1). At a point just basal to the resonant place, a peak of displacement will occur on the BM. At a point just apical to the resonant place, the BM impedance is insufficient to maintain a pressure difference across the cochlear partition, and the traveling wave dies out. The location of the displacement peak depends on stimulus frequency, so that each segment of the BM is tuned to a characteristic frequency (CF). von Békésy measured BM movement under postmortem conditions and found that the mechanical responses were broadly tuned and insensitive. Subsequent measurements made from sensitive cochleae *in vivo* have revealed that, in fact, BM displacement responses are sharply tuned and very sensitive at low sound-pressure levels (3–6).

Contributions from the electromotile outer hair cells (OHCs) (7–10) appear to be essential for cochlear sensitivity and frequency tuning (11, 12). Many current models of cochlear function include OHC-mediated feedback that amplifies BM motion at the CF to compensate for the energy dissipation that would occur through viscous damping (13–15). However, for the amplification of BM motion to take place, OHC feedback must occur at the appropriate phase of BM displacement (14, 16) and at the appropriate place along the length of the BM.

When in Each Cycle of Vibration Do OHCs Boost BM Displacement?

In a simple model of hair cell excitation (17), the BM hinges about its attachment to the spiral lamina. When the BM moves toward scala media, shearing between the tectorial membrane (TM) and the reticular lamina causes hair bundles to be deflected in the depolarizing direction toward the tallest row of stereocilia (Fig. 1a). In studies of excised cochleae (18, 19), electrical depolarization of the OHCs, which mimics the effects of excitatory hair-bundle deflection, induces somatic contractions that cause the BM and the reticular lamina to be drawn together (Fig. 1b). These findings support (14) and form the basis of (16) mechanical models of BM tuning comprising two resonant masses, e.g., represented by the BM and the reticular lamina coupled through the OHCs. However, the phase relationship between BM motion and the OHC forces must be appropriate for these models to provide frequency-dependent amplification of BM displacement. Amplification will not occur if the two resonant masses move in phase or if OHC forces are delivered in phase with the BM movement. By introducing a phase lead of 90° between the force generated by the OHCs and the passive movement of the BM, the OHCs would provide a properly timed force to enhance the BM's motion (14, 20). To deduce when in each cycle of vibration OHCs contribute to BM displacement, we investigated how different regions across the width of the BM move in response to tones at, below, and above the CF. Measurements were also made postmortem when OHCs make no active contribution to the mechanical properties of the BM (3, 11, 21).

Amplification of BM Vibrations Is Greatest in the Organ of Corti Region

The basal turn of the cochlea was exposed in guinea pigs that were under deep neuroleptanalgesia, and an opening was made in the scala tympani (22) (Fig. 1c). Through this opening, the beam of a displacement-sensitive laser-diode interferometer was focused onto the BM, forming a spot 5 μm in diameter. Tone-evoked BM displacements can be measured through the interaction of light within the laser diode and reflections from the surface of the BM. Measurements at up to 15 different

*To whom reprint requests should be addressed. E-mail: i.j.russell@sussex.ac.uk.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: BM, basilar membrane; TM, tectorial membrane; IHC, inner hair cell; OHC, outer hair cell; CF, characteristic frequency; SPL, sound pressure level; SLAM, spiral lamina attachment; SLIG, spiral ligament attachment.

locations across the width of the BM were made from nine animals in response to the CF of the recording site (22). Tones of varying sound pressure level (SPL), expressed in dB (dB relative to 2×10^{-5} Pa), were used in the experiments. The position of each measurement location was determined with respect to the outer edge of the spiral lamina ($0 \mu\text{m}$ being where the edge of the BM joins the spiral lamina). These locations were related later to the structural elements of the cochlear partition in histological sections of the measurement site (Fig. 1*d*). Locations of particular interest are indicated with horizontal bars in Fig. 1*d* and are used in subsequent figures as a reference. We determined the CF of the measurement location by constructing displacement-level functions for different frequency tones (Fig. 2*a*) and by deriving an iso-response frequency tuning curve (Fig. 2*b*). For frequencies close to the CF, displacement-level functions are typically saturating curves with steeply sloping regions at low levels and strongly compressive regions at high levels (3–6). The slope of the steep region is close to unity for frequencies half an octave or more below the CF and becomes progressively shallower as the frequencies approach and exceed the CF. This nonlinear behavior of the displacement-level functions is attributed to feedback from the OHCs, which increases as the tone frequency approaches the CF (2).

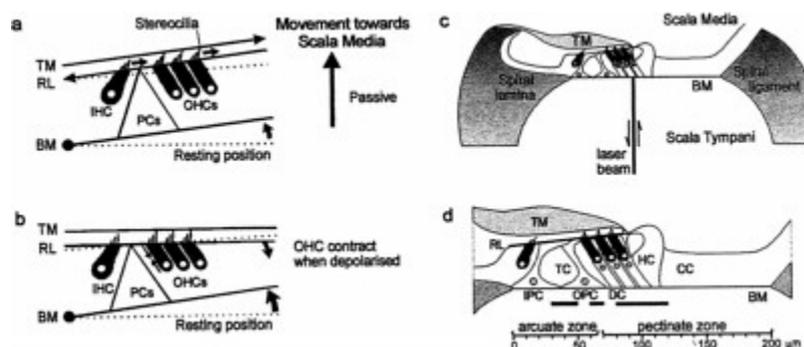


Fig. 1. (a) A model of hair cell excitation without OHC motility according to Davis (17) (see text for details), (b) OHCs contract when depolarized and the BM is drawn to the reticular lamina, (c) Experimental arrangement with laser diode beam, (d) Transverse section through the organ of Corti in the 15.5-kHz region based on measurements made *in vivo* and from histological sections: scale bar referenced to the bony edge of the spiral lamina. Solid horizontal bars indicate the following regions across the BM width with respect to the bony edge of the spiral lamina: 30–50 μm (junction of inner and outer pillar cells), 60–70 μm (near midpoint of OPC base), 80–120 μm (Deiters' cells). TM, tectorial membrane; IHC, inner hair cell; OHC, outer hair cell; HC, Hensen cell; CC, Claudius cell region; PCs, pillar cells; IPC, inner pillar cell; OPC, outer pillar cell; DC, Deiters' cell; RL, reticular lamina; TC, tunnel of Corti. Experiments were performed on deeply anaesthetized pigmented guinea pigs (180–300g, 0.06mg atropine sulfate s.c., 30mg/kg pentobarbitone i.p., 4mg/kg Droperidol i.m.; 1mg/kg Phenoperidine i.m.), which were tracheotomized, artificially respired, and with core temperatures maintained at 37°C. Modified from Nilsen and Russell (22).

The differences in displacement-level functions measured across the width of the BM (Fig. 3 *a-c*) are revealed in iso-level gain profiles (Fig. 3 *d-f*). The gain of BM displacement is calculated by dividing the displacement magnitude by the sound pressure at which the measurements are made. The degree of compression is reflected in the separation between the gain profiles at successive levels, so that the greater the separation the stronger the compression. It can be seen that the greatest compression occurs in response to frequencies at and above the CF. In all three profiles (Fig. 3 *d-f*), the greatest gain occurs beneath the organ of Corti. In response to CF tones, there are peaks of gain at the junction between the feet of the inner and outer pillar cells and in the Deiters' cell region (indicated by horizontal bars in Fig. 1*d*). These peak regions are separated by a node of minimum gain near the outer pillar cell foot (indicated by the middle bar in Fig. 1*d*). These characteristics are not clearly apparent in the gain profiles obtained at 12.5 kHz and 17.5 kHz. Displacement measurements were also made on the spiral lamina and the spiral ligament themselves, at locations that were 20 μm from the edges of the BM in a postmortem preparation (Fig. 4*e*). Vibrations were not detected above the measurement noise floor of 0.5 nm for CF tones below 90 dB SPL. The gains calculated from measurements made above 90 dB SPL (0.7 for the spiral lamina and 1.2 for the spiral ligament, at 95 dB SPL) are plotted as stars in Fig. 3 *d-f*.

Different Regions Across the BM Width Respond at Different Phases to CF Tones

To see how different regions across the width of the BM move in relation to each other, we measured the phase of BM displacement relative to the stimulus tone as a function of both measurement location and sound level. For frequencies below the CF (Fig. 3*g*), responses to low-level tones typically (3–6)

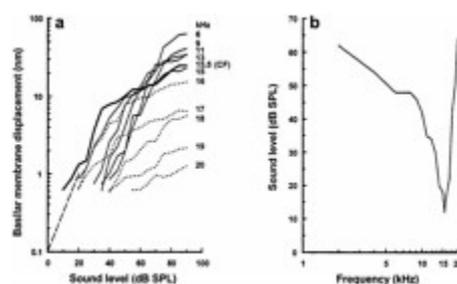


Fig. 2. (a) BM displacement as a function of sound level for frequencies between 6–20 kHz recorded from the 15.5-kHz place. Solid lines, below CF; thick line, at CF; dotted lines, above CF. Dashed line indicates slope of 1. (b) Iso-response tuning curve derived from the displacement-level functions in *a*. Response criterion: 0.7 nm [BM displacement at the detection threshold of the compound action potential (CAP) for a 15.5 kHz tone. Recording location: 100 μm from the spiral lamina. From Nilsen and Russell (22), with permission from *Nature Neuroscience* (Copyright 1999).

phase lead those to high-level tones. At the CF and above (Fig. 3 *h* and *i*), the responses of low-level tones phase lag those of high-level tones. However, measurements within about 40 μm of the spiral lamina are relatively independent of level for frequencies at and above CF, whereas those about 60–100 μm from the spiral lamina (close to the OHCs) appear to be particularly sensitive to sound level. Signal-to-noise levels in phase measurements made at frequencies above the CF and close to the spiral ligament are low, and this makes it difficult, and sometimes impossible (Fig. 3 *l*), to determine the phase of BM displacement in this region. The phase-level functions shown in Fig. 3 *g-i* are reexpressed in Fig. 3 *j-l* to show the variation of phase with distance from the spiral lamina. In response to 12.5 kHz tones, there is no clear variation in phase with location across the width of the BM, only a systematic phase lead of about 45° from the spiral lamina to the spiral ligament. However, phase varies strikingly with location for CF tones. The BM near the attachment to the spiral lamina, beneath the outer pillar cell foot and close to the spiral ligament, have similar phases and lag the BM regions beneath the inner pillar cell foot and Deiters' cells by up to 135°. Above CF, the phases of BM responses measured close to the inner pillar cells are similar to those measured at the CF in that they are relatively independent of level (Fig. 3 *l*). However, in the region around the outer pillar cell and OHC/Deiters' cell complex, the phase is highly level dependent.

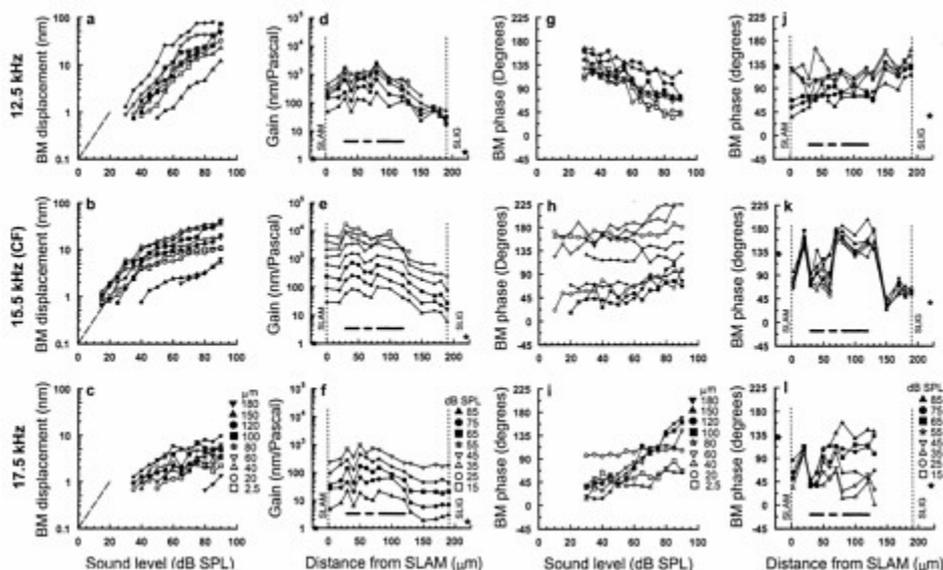


Fig. 3. (*a-c*) BM displacement-level functions in response to 12.5 kHz, 15.5 kHz (CF), and 17.5 kHz tones for different locations across the width of the BM. The locations are measured from the edge of the spiral lamina (symbols shown in *c*). Dashed line indicates slope of 1. (*d-f*) BM gain (nanometers/Pascal) in response (symbols shown in *f*). The solid stars represent BM gain of vibrations measured on the spiral lamina attachment (SLAM) (0.7) and spiral ligament attachment (SLIG) (1.7) at 95 dB SPL from a postmortem preparation, (*g-i*) BM phase as a function of the stimulus level of 12.5-kHz, 15.5-kHz (CF), and 17.5-kHz tones for different measuring locations across the width of the BM (symbols shown in *i*). (*j-l*) BM phase in response to 12.5-kHz, 15.5-kHz (CF), and 17.5-kHz tones as a function of measurement location across the width of the BM for stimulus levels between 15–85 dB SPL (symbols shown in *l*). The solid stars represent BM phase of vibrations measured on the SLAM and SLIG at 95 dB SPL from a postmortem preparation. The vertical dashed lines represent the inner SLAM and outer SLIG of the BM. Horizontal bars as in Fig. 1*d*; BM width: 190 μm ; CAP threshold at CF: 12 dB SPL; 0 dB loss after opening cochlea to expose the BM; measurement noise floor 0.7 nm. *b, e, h, and k* are from Nilsen and Russell (22), with permission from Nature Neuroscience (Copyright 1999).

Phase measurements made from the spiral lamina and spiral ligament are represented as stars in Fig. 3*j-l* and are measured from the postmortem preparation (Fig. 4*j*). Vibrations measured from the spiral ligament are in phase with those of the adjacent BM, an indication that these vibrations are driven by the BM or by pressure differences between the scala tympani and scala media. Vibrations measured from the spiral lamina have phases different from that of the adjacent BM but are the same for all frequencies, an indication that these vibrations are caused by conduction through the bones of the skull or fluids of the cochlea.

Thus, for frequencies at the CF and above, the pattern of vibration across the BM width has several modes but a single mode for frequencies below the CF.

Cochlear Sensitivity and Ambient Noise Influence BM Displacement Measurements

In Fig. 4, we show four examples of variations in gain and phase with location across the BM width that were measured under different recording conditions in response to CF tones. These figures reveal that BM displacement measurements can be influenced by 10- to 20-dB losses in sensitivity during exposure of the BM and by ambient noise. By comparison with the laser Doppler velocimeters that are commonly used to make BM measurements (4–6), BM displacements measured with a self-mixing effect interferometer are vulnerable to low-frequency vibrations, such as those caused by building noise, heartbeat, and respiration, which can be difficult to minimize. Thus at best, the measurement noise floor in our experiments is about 0.4–0.8

nm, whereas the noise floor of laser Doppler measurements is typically two orders of magnitude less than this (4–6). The data in Fig. 3 and Fig. 4*a* and *b* are from sensitive cochleae that were prepared without a perceptible loss in sensitivity and that were obtained under relatively stable recording conditions. The data in Fig. 4*c* were also collected under similar recording conditions, but a 10-dB loss in sensitivity was incurred during initial preparation. The data shown in Fig. 4*d* were obtained after a 15- to 20-dB sensitivity loss during initial preparation. In preparations where the loss in sensitivity is <10 dB, the gain and phase profiles across the width of the BM are distinctive and similar to those shown in Fig. 3 for CF tones. In the least sensitive of the preparations shown in Fig. 4*d*, BM displacement gain is still greatest over the organ of Corti but without a distinct minimum in the region of the outer pillar cell foot. The phase pattern is also disrupted in that the point of flexion near the junction between the inner and outer pillar cells appears to have moved closer to the spiral lamina, and the outer pillar cell and the OHC/Deiters' cell complex move together. However, the inner pillar cell and the OHC/Deiters' cell complex still phase lead the flexion point and the spiral ligament by up to 90°. Postmortem, when the contribution of OHC feedback is absent, responses to CF tones were detected only above the 0.5-nm noise floor when the sound level was increased above 65 dB SPL. The displacement-level functions are linear and without compression regardless of the measurement location and frequency (Fig. 4*e*). Across the BM width, the largest gains are measured beneath the outer pillar cell feet (Fig. 4*e*; note scale change) and thus resemble responses that have been recorded from insensitive living preparations (4). The phase of BM vibrations varies progressively across the BM width, with vibrations measured close to the spiral ligament leading those near the spiral lamina by about 45° (Fig. 4*j*). This variation of phase across the BM width is similar to that measured from sensitive cochleae for frequencies below the CF (Fig. 3 *g* and *j*).

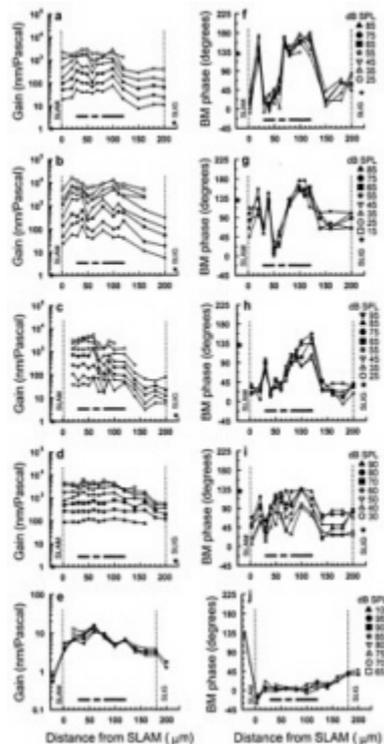


Fig. 4. BM responses as a function of measurement location across the width of the BM at different stimulus levels (symbols for each preparation shown on the right), measured under different recording conditions. (*a* and *b*) BM gain (nanometers/Pascal) and (*f* and *g*) BM phase measured from sensitive preparations with no perceptible loss in sensitivity, CF=15.5 kHz (*a* and *f*), 18 kHz (*b* and *g*). (*c*) BM gain and (*h*) BM phase measured from a preparation with a 10-dB loss of sensitivity in initial preparation, CF=16 kHz. (*d*) BM gain and (*i*) BM phase measured from a preparation with a 15- to 20-dB loss of sensitivity in initial preparation, CF=18.25 kHz. The solid stars represent BM gain and phase of vibrations measured on the SLAM (0.7) and SLIG (1.7) at 95 dB SPL from the postmortem preparation in *e*. (*e*) BM gain and (*j*) BM phase measurements made 1.5–4 h postmortem near the 15.5-kHz location of the BM. The vertical dashed lines represent the inner spiral lamina attachment (SLAM) and outer spiral ligament attachment (SLIG) of the BM. Horizontal bars as in Fig. 1*d*.

Within the limitations of our measurements, vibration patterns across the BM width in response to CF tones are similar in sensitive cochleae, and any differences can be attributed largely to variations in preparation sensitivity and ambient noise levels.

Visualizing the Relative Movement of Different Regions Across the Width of the BM

We visualized the way different regions across the width of the BM vibrate in relation to one another by plotting the product of the magnitude and cosine of the phase angle (measured relative to the sound source) of the displacement. This provides the instantaneous transverse vector of BM displacement as a function of location across the width of the BM. The solid and dotted lines in Fig. 5*a* plot $[D \cdot \cos(\theta(t))]$ at 45° intervals across the BM width for 15.5 kHz, CF tones at 40 dB SPL, where D is the magnitude of BM displacement and θ is the phase. The gray shaded area plots the envelope of the BM displacement and is widest, and BM displacement is therefore greatest, in the region of the organ of Corti. As can be seen in Fig. 3, the envelope in this region is bilobed, with peaks of displacement at the junction between the inner and outer pillar cell feet and beneath the Deiters' cells. These are separated by a minimum displacement near the center of the outer pillar cell base. Thus, through interaction with other elements in the organ of Corti, OHCs appear to limit the movement of the outer pillar cells while permitting Deiters' cells and the junction between the pillar cell feet to move more freely with respect to each other.

OHCs Boost the Vibrations of the BM at the Time of Maximum BM Velocity

The arrows in Fig. 5*a* indicate the instantaneous direction of BM motion and show that as the regions beneath the Deiters' cells and inner pillar cells move farthest into the scala media or scala tympani, the region adjacent to the spiral ligament, which is driven remotely by the OHCs, crosses its resting position and is thus moving with maximum velocity (Fig. 5*a*). In Fig. 5 *b* and *c*, these observations are interpreted in terms of OHC contraction (7–10) when the OHCs are depolarized by displacements toward scala media (23) and elongation when the OHCs are hyperpolarized.

larized by displacements toward scala tympani. Accordingly, we propose that OHCs boost the vibrations of the BM by achieving their maximum length changes and changes in force generation (24, 25) at the time of maximum BM velocity (14–16).

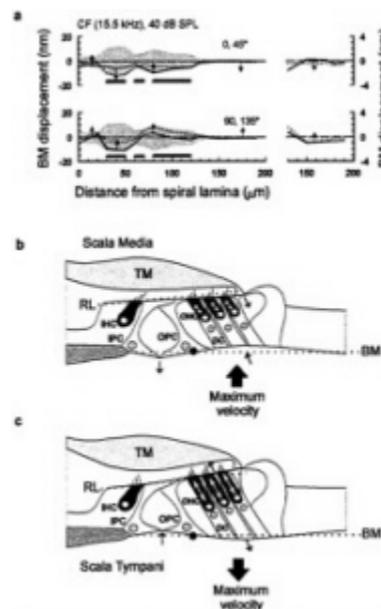


Fig. 5. (a) Displacement across the width of the BM as a function of time for a 40-dB SPL CF tone (based on data from Fig. 3). The gray shaded area plots the envelope of the BM displacement. The solid and dotted lines plot the BM displacement at successive 45° intervals, i.e., 0°, solid; 45°, dotted. The arrows indicate the instantaneous direction of BM motion in the transverse plane; horizontal bars as in Fig. 1d. (Right) Expanded view of the region spanning from 130 µm from the spiral lamina to the edge of the spiral ligament. Modified from Nilsen and Russell (22). (b and c) A simple model of BM vibration. OHCs are proposed to generate positive (cell-body shortening) forces during maximum BM velocity toward scala media (b) and negative (cell body lengthening) forces during maximum BM velocity toward scala tympani (c). The model is also proposed to account for the spatial variation in the magnitude and phase of the displacements we have measured in response to CF tones across the width of the BM. Red dotted lines, resting positions of the basilar membrane (BM) and the reticular lamina (RL); red dot, base of the outer pillar (OPC) foot which acts as a point of flexion; TM, tectorial membrane; IHC, inner hair cell; IPC, inner pillar cell; OHC, outer hair cell; DC, Deiters' cell. Arrows indicate the direction of movement. See text for explanation.

We propose a simple model to account for the spatial variation in the displacement magnitude and phase that we have measured in response to CF tones across the width of the BM (Fig. 5 b and c). When the BM undergoes excitatory displacements toward scala media (Fig. 5b), the OHCs contract with a lead of 90°, thereby drawing the BM and the reticular lamina together. In this configuration, the outer pillar cell behaves as a rigid strut (26, 27) extending from a broad foot via an articulated joint. The outer pillar cell resists the contractions of the OHCs so that when the BM is pulled up toward the scala media, the foot of the outer pillar cell rotates slightly and pushes the joint between itself and the inner pillar cell downwards toward the scala tympani. The reticular lamina seesaws about a fulcrum at the apex of the tunnel of Corti to provide a lever ratio of 4:1 in favor of the OHCs. Upward swings of the proximal edge of the reticular lamina draw up the inner pillar cell, and the region of the BM immediately underlying it, toward the scala media. The proposed tilting of the reticular lamina upward (as the BM moves toward and away from the scala media) and downward (as the BM moves toward and away from the scala tympani; Fig. 5c) would tend to amplify further shear displacement between the reticular lamina and the tectorial membrane and hence amplify the angular rotation of the inner hair cell (IHC) and OHC hair bundles.

The basis of the 90° phase lead is unknown and could, in fact, be a 270° phase delay between OHC depolarization and shortening of the OHC body because of two components. A delay of 180° could be introduced if OHC length changes were driven by potentials in the extracellular spaces surrounding the OHCs rather than by intracellular potentials in the OHCs themselves (28, 29). A further 90° delay could be introduced by a TM resonance (30) tuned to a frequency half an octave below that of the BM.

In summary, the BM beneath the Deiters' cells and OHCs vibrates with a greater magnitude and phase lead of approximately 90° compared with the BM regions beneath the bases of the outer pillar cells and adjacent to the spiral ligament. This complex vibration is not seen postmortem when amplification attributable to the OHCs is lost, resulting in a reduction in the gain of BM vibration in the Deiters' cell region by about 65 dB (5, 6).

In many respects, our measurements resemble the behavior of Kolston's three-dimensional model of cochlear mechanics (31). In particular, both the model and our measurements exhibit a bilobed displacement pattern across the width of the BM. The model also predicts that the motion beneath the Deiters' cells phase leads that of the pillar cells by about 90°. However, in the model, the inner and outer pillar cells are treated as a single element. Thus, the model does not reveal the more complex phase behavior of the BM indicated by our measurements and which we have interpreted as being caused by a flexion point at the junction between the two pillar cell bases. The model also does not predict the phase lead of about 90°, which we have measured between the region of the BM beneath the Deiters' cells and that adjacent to the spiral ligament.

Comparisons with Other Measurements

Several studies investigating mechanically and electrically evoked radial movements of the BM and organ of Corti reveal that the phase and magnitude of these movements are complex and vary with measurement position (32–36).[†] In excised cochlea segments, electrically evoked organ of Corti displacements are greatest in the region of the OHCs and are smaller in the IHC region (33). Measurements of BM stiffness (36) show the BM region around the outer pillar cell foot to be stiffer than surrounding regions. In accordance with this, measurements from the excised cochlea reveal a striking decrease in movement amplitude between the inner OHC row and the pillar cells (33). Thus, there are some strong similarities between the electrically evoked responses recorded *in vitro* and those we report here for *in vivo* BM movements measured in response to CF tones.

[†]Karavtaki, K.D. & Mountain, D.C. (1998) *Abstr. 21th Midwinter Mtg. Assoc. Res. Otolaryngol.* 719; Richter, C.P., Evans, B.N., Hu, X. & Dallos, P. (1998) *Abstr. 21th Midwinter Mtg. Assoc. Res. Otolaryngol.* 720.

In vivo measurements of acoustically driven BM movements in the horseshoe bat *Rhinolophus* (32) and electrically driven BM movements in the guinea pig cochlea reveal that the arcuate and pectinate zones move 180° out of phase (35), with the outer pillar cell foot moving as much as, or more than, the OHC region (35). These latter findings (35) differ from those reported here where the phase difference between the outer pillar cell foot and the OHC region is only about 90° with CF tone stimulation, and the OHC region moves more than the outer pillar cell foot. Interestingly, Kolston's model (31) predicts these differences in the response of the cochlear partition to local electrical and acoustic stimulation. Furthermore, our proposal that the reticular lamina tilts about a fulcrum at the apex of the tunnel of Corti is supported by direct observations of just such a motion in response to electrical stimulation of the organ of Corti in an isolated preparation (37).

Recent reports of the radial pattern of BM motion evoked by CF tones in guinea pigs measured with laser Doppler interferometry (38) ‡ without using reflective beads reveal that the BM vibrates most near the feet of the outer pillar cells. In the most sensitive preparations and at low sound levels, the region of the BM between the outer pillar cell and the spiral lamina leads the region between the outer pillar cell and the spiral ligament by about 30°. In a recent report (39) where multiple large (25 µm) reflective beads were attached to the BM near the base of the chinchilla cochlea, it was reported that there is little variation in the phase of BM motion radially across the width of the BM. Thus, although there is very good agreement between single point measurements on the BM using self-mixing interferometry and laser Doppler interferometry and indeed for measurements along the length of the BM (see below), there are significant differences between measurements that have been made across the width of the BM.§

Where Along the Traveling Wave Is Energy Introduced?

A number of modeling studies have suggested that energy is introduced at a point in the BM traveling wave basal to the CF place and that the active element extends over several millimeters (11, 43, 44). This proposal is attractive because the traveling wave would pass through the active region before reaching the CF, thereby directly stimulating the OHCs, which could then feed energy back to boost the displacement of the adjacent CF region. The location and extent of the BM region responsible for amplifying the traveling wave can be obtained by measuring the spatial representation of the nonlinear responses to a pure tone along the BM. Spatial representations of pure tones on the BM have been constructed from the responses of auditory afferent nerve fibers in sensitive cochleae. However, these indirect measurements of excitation pattern were confined to low frequencies (45–47). We have measured the distribution along the BM of nonlinear compressive vibrations to 15- and 15.5-kHz tones in the 12.5- to 27-kHz region of the guinea pig cochlea *in vivo*.

Self-mixing interferometry has considerable drawbacks for making spatial measurements. It is essential to focus and maximize the signal in terms of both sensitivity and intensity at each measurement location. If these conditions are obeyed, it is possible to plot a surface profile of the BM. If the beam is not focused, it is still reflected from this surface but from a broad diffuse spot. It is essential to adhere to these requirements and to check for repeatability of measurement. Therefore, data acquisition is slow. It is necessary to repeat all measurements if the recording conditions change and to abandon measurements if ambient noise cannot be controlled or if the sensitivity of the animal decreases by more than 5 dB during the measurements. Thus measurements are slow to perform, and our success rate is low.

BM Displacements in Response to a Tone Are Amplified Most at the Tone's Frequency Place

In response to a tone (15 kHz, Fig. 6a; 15.5 kHz, Fig. 6b), the gains of displacement were measured at up to 15 different locations along the length of the BM for sound levels 15–100 dB SPL. The exact location of each measurement point with respect to the apex of the cochlea was determined by using lesion studies (48) and a function relating the CF of a point along the BM to its distance from the apex (49). At each recording site, we determined the CF and used this to estimate the distance of the measurement point from the apex of the cochlea. Thus, the spatial distribution of the tone was obtained along the length of the BM. This was achieved for a varying number of measurement sites in six preparations. For low levels of the tone, the maximum gain and spread of excitation are centered on the tone's frequency place (Fig. 6 a and b, dashed lines). The gain of BM displacement declines progressively with sound level at the frequency place, being 1,000 times greater at 15 than at 100 dB SPL, when presumably the vibrations of the BM are dominated by the passive properties of the cochlear partition. The gain also decreases rapidly with distance from the frequency place. There is a sharp cutoff in the apical extent of the response to the tone near the 12.5-kHz place, when the acoustic impedance of the BM becomes dominated by its mass and the BM appears unable to sustain either pressure differences between the scalae vestibuli and tympani or any OHC forces that attempt to boost its motion at 15.5 kHz. A decline in gain also occurs with distance in the basal direction, so that at about 0.7 mm from the frequency place of the tone, the gain is reduced to about 10% of its value at the frequency place. This is presumably because OHC forces generated at the 15.5-kHz place are ineffective against the stiffness-dominated acoustic impedance of the BM at locations basal to the 15.5-kHz place.

The "Panoramic" Tuning Curve Peak Shifts Toward the Cochlea Base at High SPLs

According to the measurements shown in Fig. 6a, the peak response to the 15-kHz tone remains almost stationary in its location on the BM for levels below about 80 dB SPL. The peak response from another preparation shown in Fig. 6b moves toward the base of the cochlea when the level exceeds 55 dB SPL, in accord with the shift at high levels in the tuning curve peak of cochlear responses to a frequency about half an octave below the CF (50, 51). Differences between the two preparations in the SPL threshold for the basal-ward shift in the peak are not caused by differences in sensitivity but represent variability between preparations that may be caused by differences in their susceptibility to intense tones. Differences between preparations in the threshold for the half-octave shift, although not as marked as

‡Cooper, N.P. (2000) *Abstr. 23rd Midwinter Mtg. Assoc. Res. Otolaryngol.* 254.

§Displacement-sensitive self-mixing interferometry does not require the placement of reflective beads on the BM, and we can accurately choose the sites of measurement. There is strong disagreement as to whether beads either remain firmly attached to the BM or are influenced by hydrodynamic forces resulting from movements of the BM (40, 41). A striking example that hydrodynamic forces at the organ of Corti can be quite predominant is seen in the responses of hair cells in the cochleae of mice without attached tectorial membranes that can respond to BM displacement as a consequence of viscous drag acting on their hair bundles [Russell, I.J., Lukashkina, V.A., Kössl, M., Legan, K., Goodyear, R. & Richardson, G.P. (2000) *Abstr. 23rd Midwinter Mtg. Assoc. Res. Otolaryngol.* 249]. We are not sure whether the Doppler shift in frequency of the laser beam, on which BM velocity and some displacement measurements are based, necessarily reflects movements in the transverse plane of BM motion. We were concerned with this last aspect of measurement when we considered using the Doppler frequency shift that results when light is coupled back into the laser diode, and which is widely used to measure velocity in the plane at right angles to the beam (42). In view of this, we determined early on (unpublished work) that any transverse movements of the BM that might be caused by 10- to 30-nm vibrations of the preparation in the radial plane were not detectable above the 0.4- to 0.8-nm noise floor of our measurement system.

those in Fig. 6 *a* and *b*, may be observed in other preparations (6, 11, 39).

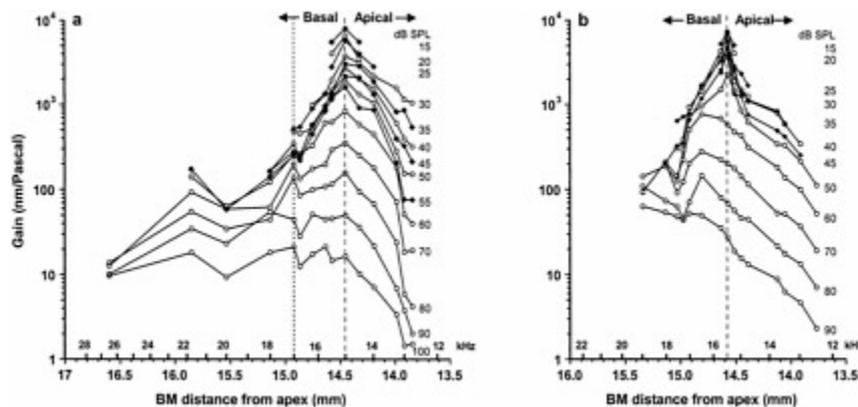


Fig. 6. The gain of BM displacements in response to 15-kHz (*a*) and 15.5-kHz tones (*b*) measured at 15 positions along the BM over a range of intensities from 15 to 100 dB SPL. Data from *a* and *b* are from separate preparations. Dashed line indicates the CF location. In *a*, frequency positions to the right of the dotted line (12.5–17 kHz) are from one preparation and those on the left (17–27 kHz) are each from separate preparations. Data in *a* modified from Nilsen and Russell (48).

BM Phase Measurements Indicate That the CF Place Is the Point of Resonance for a CF Tone

The phase of displacement evoked by responses to low-level 15 kHz tones is plotted relative to that of a higher-level (70 dB SPL) 15-kHz tone (Fig. 7*a*) for the preparation shown in Fig. 6*a*. It can be seen that at the 15-kHz place, the phase is relatively independent of sound level. For locations just basal to the 15-kHz place, the low-level responses lead those of the high-level responses. For locations apical to the 15-kHz place, the phase of low-level responses strongly lag those of the high-level responses. Accordingly, the 15-kHz place behaves as a point of resonance on the BM. Observations from this preparation and others show that in response to 15-kHz tones, level-dependent phase changes are not evident in BM regions more basal than the 17-kHz place. This may indicate that feedback from OHCs located basal to the 17-kHz location is unlikely to contribute to boosting BM responses to 15-kHz tones. Similar findings were obtained for the preparation shown in Fig. 6*b* for 15.5-kHz tones.

OHC Feedback Is Restricted to a Limited Region Extending Apical and Basal to the CF Site

In a linear system, the slopes of input-output functions are 1 dB/dB. In systems with feedback that changes as a function of stimulus level, response-level functions can grow at less than 1 dB/dB. This is the case in the cochlea where for CF tones and for sound levels just above the response threshold of auditory fibers the slopes of BM displacement-level functions lie between 0.8 and 0.5 dB/dB. For sound levels above about 40 dB SPL, the displacement-level functions become strongly compressive with slopes of about 0.2 dB/dB. We found that for locations between the 12.5-kHz and 16.5-kHz places on the BM, 15-kHz displacement-level functions were nonlinear and compressive and, therefore, were subject to nonlinear amplification from the OHCs. The slopes and magnitude at which the displacement saturated decreased as the measuring point was moved apical to the 15-kHz point and toward lower frequencies (Fig. 7*b*). At locations basal to the 16.5-kHz place, where amplification is greatly reduced, the steep slopes of the displacement-level functions become unity (Fig. 7*c*). We thus suggest that the 1.25-mm long section of the BM between the 12.5- to 16.5-kHz regions is the source of mechanically active elements contributing to the amplification of the 15-kHz place. On a similar basis, we estimated that a 1.4-mm-long section of the BM contributed to the 15.5-kHz place. The extent and location of the nonlinear region of BM displacement reported here is in agreement with estimates of the distribution of the cochlear amplifier based on noise-induced lesions of the BM (52). Our estimate is somewhat less than that of the extent of the cochlear amplifier in the basal turn of the guinea pig cochlea based on power dissipation and modeling of the cochlea (53). It is also less than the 2.0- to 3.5-mm estimates of the cochlear amplifier based on the spatial

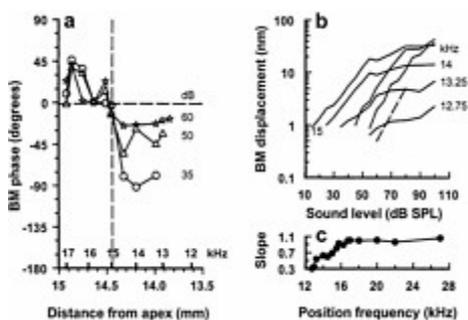


Fig. 7. (*a*) The phase of BM displacement evoked by responses to low-level 15-kHz tones plotted relative to that of a higher-level (70 dB SPL) 15-kHz tone at different positions along the BM. Vertical dashed line indicates 15-kHz (CF) location. (*b*) BM displacement-level functions in response to a 15-kHz tone measured over a range of frequency positions about the 15-kHz point. Dashed line indicates a slope of 1. (*c*) The slope of the initial region of the displacement-level function as a function of frequency position. Modified from Russell and Nilsen (48).

extent of the compressive region in the basal region of the chinchilla cochlea (39).

Conclusions

In conclusion, we propose that OHCs boost the vibrations of the BM by achieving their maximum length changes and changes in force generation (24, 25) at the time of maximum BM velocity (14, 16). We find that compressive nonlinear displacements of the BM in response to 15- and 15.5-kHz tones are restricted to a 1.25- to 1.40-mm region extending both apical and basal to the frequency place of the tone on the BM. If this nonlinearity reflects OHC-mediated feedback to the cochlear partition, no more than about 500 OHCs (48) contribute to the displacement of the 15- and 15.5-kHz frequency places on the BM.

We thank Manfred Kössl, Andrei Lukashkin, and Guy Richardson for constructive comments on an early draft of the paper. This work was supported by grants from the Medical Research Council and Wellcome Trust.

1. von Békésy, G. (1960) in *Experiments in Hearing* (McGraw-Hill, New York).
2. Geisler, C.D. (1998) in *From Sound to Synapse: Physiology of the Mammalian Ear* (Oxford Univ. Press, Oxford), 60–64.
3. Sellick, P.M., Patuzzi, R. & Johnstone, B.M. (1982) *J. Acoust. Soc. Am.* **72**, 131–141.
4. Cooper, N.P. & Rhode, W.S. (1992) *Hear. Res.* **63**, 163–190.
5. Nuttall, A.L. & Dolan, D.F. (1996) *J. Acoust. Soc. Am.* **99**, 1556–1565.
6. Ruggero, M.A., Rich, N.C., Recio, A., Shyamla Narayan, S. & Robles, L. (1997) *J. Acoust. Soc. Am.* **101**, 2151–2163.
7. Brownell, W.E., Bader, C.R., Bertrand, D. & de Ribaupierre, Y. (1985) *Science* **227**, 194–196.
8. Ashmore, J.F. (1987) *J. Physiol.* **388**, 323–347.
9. Dallos, P., Evans, B.N. & Hallworth, R. (1991) *Nature (London)* **350**, 155–157.
10. Santos-Sacchi, J. (1992) *J. Neurosci.* **12**, 1906–1916.
11. Ruggero, M.A. & Rich, N.C. (1991) *J. Neurosci.* **11**, 1057–1067.
12. Murugasu, E. & Russell, I.J. (1996) *J. Neurosci.* **16**, 325–332.
13. Neely S.T. & Kim, D.O. (1986) *J. Acoust. Soc. Am.* **79**, 567–573.
14. Geisler, C.D. & Sang, C.N. (1995) *Hear. Res.* **86**, 132–146.
15. Nobili, R. & Mammano, F. (1996) *J. Acoust. Soc. Am.* **99**, 2244–2255.
16. Markin, V.S. & Hudspeth, A.J. (1995) *Biophys. J.* **69**, 138–147.
17. Davis, H. (1958) *Laryngoscope* **68**, 359–382.
18. Mammano, F. & Ashmore, J.F. (1993) *Nature (London)* **365**, 838–841.
19. Mammano, F., Kros, C.J. & Ashmore, J.F. (1995) *Pflügers Arch.* **430**, 745–750.
20. Robles, L., Ruggero, M.A. & Rich, N.C. (1986) *J. Acoust. Soc. Am.* **80**, 1364–1374.
21. Nuttall, A.L., Dolan, D.F. & Avinash, G. (1991) *Hear. Res.* **51**, 203–214.
22. Nilsen, K.E. & Russell, I.J. (1999) *Nat. Neurosci.* **2**, 642–648.
23. Russell, I.J. & Sellick, P.M. (1983) *J. Physiol. (London)* **338**, 179–206.
24. Hallworth, R. (1995) *J. Neurophysiol.* **74**, 2319–2328.
25. Frank, G., Hemmert, W. & Gummer, A.W. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4420–4425.
26. Tolomeo, J.A. & Holley, M.C. (1997) *Biophys. J.* **73**, 2241–2247.
27. Mogensen, M.M., Henderson, C.G., Mackie, J.B., Lane, E.B., Garrod, D.R. & Tucker, J.B. (1998) *Cell Motil. Cytoskeleton* **41**, 138–153.
28. Kolston, P.J. (1995) *Trends Neurosci.* **18**, 427–429.
29. Dallos, P. & Evans, B.N. (1995) *Science* **267**, 2006–2009.
30. Gummer, A.W., Hemmert, W. & Zenner, H-P. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 8727–8732.
31. Kolston, P.J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 3676–3681.
32. Wilson, J.P. & Bruns, V. (1983) *Hear. Res.* **10**, 15–35.
33. Reuter, G. & Zenner, H. (1990) *Hear. Res.* **43**, 219–230.
34. Xue, S., Mountain, D.C. & Hubbard, A.E. (1993) in *Biophysics of Hair Cell Sensory Systems*, eds. Duifhuis, H., Horst, J.W., van Dijk, P. & Netten, S.M. (World Scientific, Singapore), 370–376.
35. Nuttall, A.L., Guo, M.H. & Ren, T.Y. (1999) *Hearing Res.* **131**, 39–46.
36. Olson, E.S. & Mountain, D.C. (1994) *J. Acoust. Soc. Am.* **95**, 395–400.
37. Frank, G., Scherer, M., Hemmert, W., Zenner, H-P. & Gummer, A.W. (1999) in *Recent Developments in Auditory Mechanics*, eds. Wada, H. & Takasaka, T. (World Scientific, Singapore).
38. Cooper, N.P. (1999) in *Recent Developments in Auditory Mechanics*, eds. Wada, H. & Takasaka, T. (World Scientific, Singapore).
39. Rhode, W.S. & Reccio, A. (2000) *J. Acoust. Soc. Am.* **107**, 3317–3332.
40. Khanna, S.M., Ulfendahl, M. & Steele, C.R. (1988) *Hearing Res.* **116**, 71–85.
41. Cooper, N. (1999) *J. Acoust. Soc. Am.* **106**, L59–L64.
42. Koelink, M.H., Slot, M., de Mul, F.F.M., Greve, J., Graaff, R., Dassel, A.C.M. & Aardnoudse, J.G. (1992) *Appl. Opt.* **31**, 3401–3408.
43. de Boer, E. (1983) *J. Acoust. Soc. Am.* **73**, 567–573.
44. Diependaal, R.J., de Boer, E. & Viergever, M.A. (1987) *J. Acoust. Soc. Am.* **80**, 124–132.
45. Pfeiffer, R.R. & Kim, D.O. (1975) *J. Acoust. Soc. Am.* **58**, 867–869.
46. Kim, D.O. & Molnar, C.E. (1979) *J. Neurophysiol.* **42**, 16–30.
47. Geisler, C.D. & Cai, Y. (1996) *J. Acoust. Soc. Am.* **99**, 1550–1555.
48. Russell, I.J. & Nilsen, K.E. (1997) *Proc. Nat. Acad. Sci. USA* **94**, 2660–2664.
49. Greenwood, D.D. (1990) *J. Acoust. Soc. Am.* **87**, 2592–2605.
50. Mitchell, C., Brummet, R.E. & Vernon, J.A. (1977) *Arch. Otolaryngol.* **103**, 117–123.
51. Cody, A.R. & Johnstone, B.M. (1980) *Hearing Res.* **3**, 3–16.
52. Cody, A.R. (1992) *Hearing Res.* **62**, 166–171.
53. De Boer, E. & Nuttall, A.F. (2000) *J. Acoust. Soc. Am.* **107**, 1497–1507.

MOLECULAR MECHANISMS OF SOUND AMPLIFICATION IN THE MAMMALIAN COCHLEA

Jonathan F. Ashmore*, Gwénaëlle S.G. Géléoc†, and Lene Harbott

Department of Physiology, University College London, Gower Street, London WC1E 6BT, United Kingdom

Mammalian hearing depends on the enhanced mechanical properties of the basilar membrane within the cochlear duct. The enhancement arises through the action of outer hair cells that act like force generators within the organ of Corti. Simple considerations show that underlying mechanism of somatic motility depends on local area changes within the lateral membrane of the cell. The molecular basis for this phenomenon is a dense array of particles that are inserted into the basolateral membrane and that are capable of sensing membrane potential field. We show here that outer hair cells selectively take up fructose, at rates high enough to suggest that a sugar transporter may be part of the motor complex. The relation of these findings to a recent candidate for the molecular motor is also discussed.

The inner ear of mammals has evolved to analyze sounds over a wide range of frequencies. In humans this range covers about 8 octaves, but it can be both more restricted (hearing in mice covers only 3.5 octaves) or, as in some cetaceans, extend over 10 octaves and use infrasound. Auditory specialists such as the echolocating bats may employ frequencies well into the ultrasonic. When considering how a hearing organ of the size of a pea or even smaller can perform such engineering feats, cochlear construction provides the clues. The cochlea in all these diverse mammalian species shows a remarkable conservation of design. A common feature of all mammalian cochleae is that, within the duct, the basilar membrane supports a propagated traveling wave. In mammals, the basilar membrane performs as a mechanical selector of sound frequencies and maps component frequencies in a complex sound onto a position in the duct. A second feature found throughout mammalian cochleae is the presence of two morphologically distinct sets of hair cells, inner (IHCs) and outer hair cells (OHCs). The cells lie within the organ of Corti, which runs along the full length of the basilar membrane and is shown in cross section in Fig. 1a. Hence the linkage between basilar membrane motion and the deflection of the sensory hair bundles of hair cells recodes sound frequency and intensity into a firing pattern of the afferents of the auditory nerve. The basilar membrane thus acts as a preconditioner of the sound signal. Other vertebrate genera use modified designs of hearing organ (using, for example, local filtering based on electrical resonance of their membranes or mechanical resonance of their hair cell stereocilia). The mammalian cochlear design ensures that even high frequencies can be detected by using material and mechanical properties of the macroscopic structures of the cochlea (1).

A wide range of measurements have now shown that the basilar membrane has a mechanical pattern of vibration that is under physiological control. Optimally, threshold sound detection elicits a peak deflection of about 0.3 nm. This exceeds what a membrane of the same mechanical construction can achieve and an energy source is necessarily involved. Indeed, it was suspected more than 50 years ago that there had to be some "active" process in the cochlea to account for the known psychophysics of the auditory periphery (2). The process by which enhanced basilar membrane mechanics is generated has been termed "cochlear amplification." How such amplification comes about has been the focus of much effort over the past two decades. Sensory hair cells are clearly implicated. There remains little doubt that inner hair cells are the primary sensory cells of the mammalian cochlea: they signal via a glutamatergic synapse to the auditory nerve. However, there are now many lines of evidence that OHCs of the cochlea are responsible for the enhancement of the basilar membrane motion, an idea that was proposed on the basis of ultrastructural evidence alone before direct physiological data became available (3). In addition, although OHCs also have an afferent innervation, the main neural pathway associated with them is a cholinergic efferent fiber system terminating on the cells. The organization is reminiscent of the innervation of an effector system. Most significantly, however, OHCs are positioned and coupled to the motion of the basilar membrane in a manner that does allow forces to be fed back into the dynamics of the cochlea.

Most computational models of peripheral hearing indicate cochlear amplification comes about by opposing dissipative forces within the cochlea. Such forces arise from the fluid content of the cochlear duct and the viscoelastic properties of the tissues of the basilar membrane. This situation places a number of constraints on how any mechanism dependent on the OHCs can operate. One proposal for the origin of OHC forces, for which there is evidence in nonmammalian species, is that the apical stereocilia act both as the sensors of the motion of the basilar membrane and as a motor source to amplify the motion (4). The uniformity of cochlear structure along the length of the duct implies that OHC parameters have to be tuned to match the mechanical impedance at each place (5).

The second proposal for how OHCs contribute to cochlear function invokes a novel cellular motor. This is the mechanism of "somatic electronic motility." The key observation is that cell hyperpolarization lengthens isolated cells, whereas depolarization leads to their shortening (6). This implies that OHCs are capable of generating forces. The mechanism is fast (7–9), and cells can be electrically responsive up to frequencies above 40 kHz. The sequence is that a shear displacement of the OHC stereocilia generate a change in membrane potential that in turn generates forces directed along the cell long axis. The process thus decouples the motor force from the shear forces (Fig. 1a). Computational models indicate that the range of frequencies presented as the input signal to any group of OHCs may be preselected by a broadly tuned mechanical resonance of the

*To whom reprint requests should be addressed. E-mail: j.ashmore@ucl.ac.uk.

†Present address: Harvard Medical School, Massachusetts General Hospital, Howard Hughes Medical Institute, Wellman 414, Boston, MA 02114.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviation: OHC, outer hair cell.

overlying tectorial membrane (10, 11). For this hypothesis there is now better experimental evidence (12). Thus during every cycle when the OHC is stimulated by displacement of its stereocilia, the OHC motor completes the force feedback loop to enhance the mechanical stimulus.

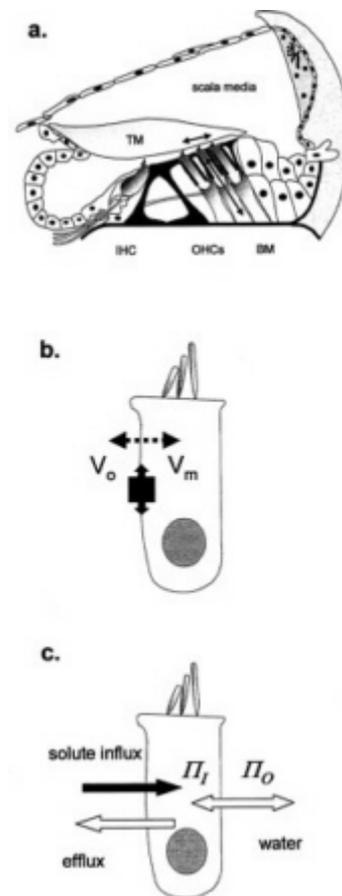


Fig. 1. OHCs in the cochlea, (a) Schematic cross section of the organ of Corti, showing site of inner hair cells (IHC) and OHCs. The primary stimulus is the shear delivered to the OHC stereocilia by the tectorial membrane (TM). OHC length changes (and therefore forces) are produced as arrowed. BM, basilar membrane. (b) OHC length change through electromotility, where membrane potential ($V_o - V_m$) alters cell surface area. The tight molecular packing in the lateral membrane allows the protein area change to have macroscopic effects, (c) OHC cell length change through cell volume change, where osmotic pressure ($\pi_i - \pi_o$) difference inside and outside requires water to follow solute entry.

The mechanism of somatic electromotility copes readily with the phasing requirement on the force operating in the cochlea. Because deflection of the stereocilia causes current to enter the cell, 90° in advance of the voltage change, the motor, being voltage sensitive, responds with the right phase to oppose viscous forces. However, at high frequencies, most of the current passes through the membrane capacitance, and it is a contentious topic as to whether the transmembrane potential across the OHC membrane is sufficient to drive the motor. Although the membrane time constant effectively filters changes in OHC membrane potential, it has been proposed that the extracellular current that flows around an OHC from neighboring cells alters the effective field across the OHC basolateral membrane (8) and thereby provides sufficient driving potential for the motor (13). This mechanism requires further modeling once the electrical parameters of the cochlea become clear.

The underlying biophysics of electromotility is constrained by the requirement that OHC mechanisms generate forces at high frequencies. This allows OHCs to be involved in the acoustic processing of sound at frequencies in the kilohertz range. The simplest of the cellular models to describe electromotility is one in which the cell membrane area changes but the volume of the cell remains constant (14). As a result, a reduction of the unit area of lateral membrane will tend to make the cell shorter. In principle this is a dynamically favored mechanism for rapid length change, as no water needs to move (cyclically) back and forth across the membrane. Although an increase of volume at constant cell surface area will also change the length of a cylindrical cell, the movement of water would be anticipated to be a slower process (Fig. 1 b and c).

The lateral membrane of OHCs contains a high density of particles that can be seen either after deep-etch and low-angle rotary shadowing (15) or by freeze fracture of the membrane P-face (16). The observed particles are about 8 nm in diameter and close packed at $5,000 \mu\text{m}^{-2}$, making up a significant fraction of the membrane. The particles are candidates for the OHC motor. The electrophysiological signature of the motor is a gating charge that can be detected both in whole cell recording (17) and by wide-band patch-clamp recording (18). The gating charge can most simply be interpreted as a dipole reorientation of some component of the motor protein within the electrical field. The gating charge is about 10 times larger than the gating charge measured in the excitation contraction coupling of muscle. As in muscle, the dipole of the gating charge gives rise to a voltage-dependent capacitance of the cell membrane (19). In OHCs the apparent maximal membrane capacitance is found near -30 mV but is tension sensitive. The presence of this motor can almost double the geometric capacitance of the cell. By matching the electrophysiological measurements with the electron microscopy, each particle, tentatively identified with the motor, appears to be associated with the movement of about 1 electronic charge across the membrane field.

It seems unlikely that the motor molecule in the OHC membrane is a modified (but necessarily nonconducting) ion channel. The voltage sensor in an ion channel, such as the S4 region, does not move within the membrane with sufficient speed. An alternative proposal is that the motor is a charged ion transporter. Cation transporters are known to be associated with rapid charge movements as part of the transport cycle and, although less well characterized, many anion transporters share this property.

A third class of candidate proteins includes those transporters that are electrically neutral. Such transporters may not reveal

themselves electrophysiologically, as no net charge is carried during their transport cycle, but there is no *a priori* reason why such proteins cannot have a gating charge movement. This class includes those proteins implicated in water or in neutral solute transport. The evidence below suggests that the OHC membrane contains a transporter with similarities to the sugar transporter, GLUT5, whose normal substrate is fructose. Sugar molecules cannot cross lipid membrane by simple diffusion. GLUT5 is a member of a family of six facilitative hexose transporters (GLUT1–7) found in mammalian tissue. Each differs in its substrate specificity and tissue distribution. Sugars entering the cell can provide the energy required by OHCs for metabolic activities, but the only such transporter so far identified by immunohistochemistry is GLUT5 (20). Although originally identified in gerbils, the same epitope is also present in guinea pig OHCs (21).

By identifying a fructose transporting mechanism in OHCs that is distributed along the length of the basolateral membrane (21), we have been led to explore the possibility that GLUT5 or a closely related isoform is also part of the motor complex. The experiments below examine quantitatively the uptake rates of different sugars into OHCs and the effect of known blockers for sugar transport and motility.

Methods

Adult guinea pigs (200–400 g) were killed by rapid cervical dislocation, and both bullae were removed. The organ of Corti was dissected in standard saline containing (in mM): NaCl, 142; KCl, 4; CaCl₂, 1; MgCl₂, 1.5; Hepes, 10; pH 7.35; osmolarity adjusted to 325 mosmol·kg⁻¹ with approximately 30 mM sucrose. The tissue was then bathed in 0.25 mg·ml⁻¹ trypsin (Sigma) for 10 min before gentle mechanical dissociation. The cells were transferred to a 1-ml chamber continuously perfused at a rate of 100 μl·min⁻¹. Cells were used within 3 h of the dissection. Experiments were performed at room temperature (20–25°C).

Sugar transport was estimated as described previously (21) by measuring OHC length. Custom software was used to estimate the cell length changes produced by water entry. Sugars were added around the cell through a puffer pipette at different concentrations (3, 10, 20, and 30 mM) in standard saline solution, adjusted to 325 mosmol·kg⁻¹ in all cases with sucrose. To compensate, 10 mM NaCl was removed from both bathing and perfusing solution in experiments where 50 mM glucose or fructose was tested. The superfusion rate from the applicator pipette was adjusted to 8–12 cm of H₂O so that movement of the cell due to pressure was minimized. Cells were observed by a 1.3 numerical aperture 40X objective and recorded at 1 Hz by a high-resolution video camera using Axon Imaging Workbench software (Axon Instruments; Foster City, CA). The same software was used for image analysis and was enhanced by pixel estimation using MATLAB 5.3 (Mathworks; Natick, MA). In digitized images the pixel diameter was 114 nm. By analysis of the centroid of the 1-μm² regions, we estimate that the errors in cell length and width measurements did not exceed 15%. Data are shown as mean ± SD and, where appropriate, were fitted to theoretical models by using a Marquadt-Levenberg algorithm.

Results

With the bath solution containing sucrose as the balancing osmolyte, solutions were applied containing isotonic glucose or fructose. With a 30-s exposure, both 30 mM fructose and 30 mM glucose produced a reversible shortening of the cell. Typical strains measured in apical cells 40–65 μm long were close to 5%, at 5.4%±0.7% (*n*=6) in fructose and 4.9%±0.7% (*n*=4) in glucose.

Isotonic replacement of sucrose with various concentrations of glucose or fructose was performed to assess substrate specificity. Fig. 2 shows that both sugars were transported into the cell in a graded and saturable manner characteristic of an uptake carrier. The affinity for fructose was higher than that for glucose, even though the maximal uptake rates were comparable; however, a significant difference between glucose- and fructose-induced strains was seen when these sugars were applied at 3 mM, 10 mM, or 20 mM. This result implies that the transport rate for glucose across the OHC membrane is less than for fructose.

The data for the initial rate of cell shortening V_i were fitted by Michaelis-Menten kinetics for saturable uptake mechanism (Fig. 2c) with V_{max} the maximum rate, $[S]$ the sugar concentration, and K_m the half-saturating concentration:

$$V_i = \frac{V_{max}[S]}{K_m + [S]}, \quad [1]$$

For fructose $K_m=15.8\pm 1.6$ mM and for glucose $K_m=34.0\pm 3.5$ mM (34 experiments, *n*=23 cells). This is typical of values reported for GLUT5 in other cell systems—for example, in human enterocytes (22). These values for K_m are related directly to sugar transport, since water equilibrated across the cell membrane at a faster rate than either of the sugars. The measured cell shortening rate obtained in hypotonic solution (switching to sucrose, 315 mosmol·kg⁻¹ from 330 mosmol·kg⁻¹) induced a faster change in length (0.71±0.13 s⁻¹) compared with 30 mM fructose [0.52±0.18 s⁻¹ (*n*=3)].

Cytochalasin B is known to act as a reversible and noncompetitive inhibitor of glucose transport by GLUT1, -2, and -3. Cytochalasin B added (at 1 mg/ml) to the bath 10 min before experiments produced no significant difference on either the uptake rates or the steady-state strains in four cells. Additionally, OHCs did not respond to isotonic replacement of external sucrose with the sugar 2-deoxy-D-glucose in 30 mM solution (Fig. 2d). This sugar is also not transported by GLUT5 isoforms (23).

Acetylsalicylic acid blocks electromotility and OHC motor charge movement (24). If the sugar uptake carrier shares properties of the OHC motor protein, we would predict that the transport itself might be affected by salicylic acid. We tested this hypothesis by replacing glucose with fructose isotonically in the presence of 10 mM salicylate (Fig. 3). Cells were pretreated with salicylate for 1 min before fructose application. As salicylate entered the cell (as salicylic acid) the cell increased in volume and shortened, consistent with salicylate loading of the cell. After a rapid phase of shortening, the cell volume stabilized. No further change in cell length was observed when fructose was applied. Application of fructose before and after washout showed that fructose-induced strain was larger and faster than strain induced by salicylate alone.

There is no evidence that significant current can pass through the OHC lateral membrane (18, 25), most of the ion channels being localized at the base or apex of the cell. The lateral cell membrane, however, exhibits a voltage-dependent capacitance, as seen in Fig. 4. The bell-shaped curve is the nonlinear capacitance

$$C_{\text{nonlinear}}(V) = C_{\text{max}} 4 \exp(-\beta(V-V_o)) / [1 + \exp(-\beta(V-V_o))]^2, \quad [2]$$

where the peak capacitance C_{max} occurs at a membrane potential V_o and the parameter β is a measure of the voltage sensitivity of the charge movement. The voltage dependence of charge movement of the membrane depended on the species of sugar presented to the OHC (Fig. 4). The largest shift in V_o at the peak membrane capacitance occurred when fructose was present around the cell. Although hypoosmotic solutions also swell the cell, inducing membrane strain and a consequent positive shift in V_o of the voltage-dependent capacitance (26, 27), the effect of

fructose was more pronounced than the osmotic stimulus that produced the same strain (Fig. 4b).

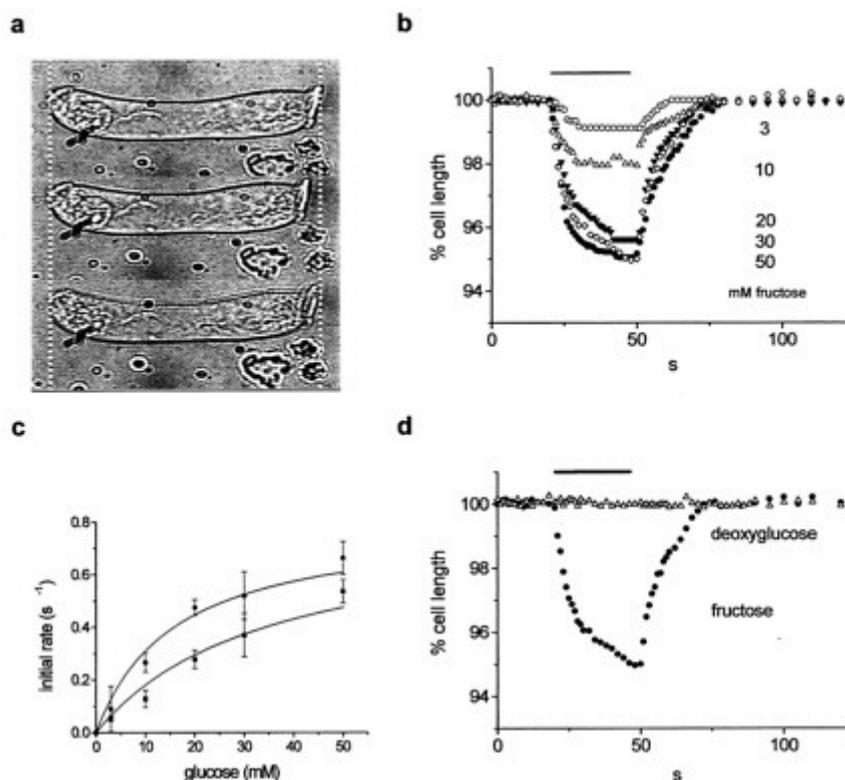


Fig. 2. Glucose and fructose uptake by guinea-pig OHCs. (a) Fructose-induced cell shortening. Top, control; middle, with fructose; bottom, washout. Cell length, 55 μm . (b) Strains at these different concentrations. Sugars were applied at 3 ($n=6$), 10 ($n=7$), 20 ($n=8$), 30 ($n=6$), and 50 mM ($n=7$). (c) Initial transport rate estimated by a linear fit between $t=0$ s and $t=6$ s of glucose (\square) or fructose (\circ) application. Data were fitted with $V_{\text{max}}=0.8\% \text{ s}^{-1}$ and $K_m=15.8\pm 1.6$ mM (fructose) and 34 ± 3.5 mM (glucose), (d) Effect of 30 mM deoxyglucose (Δ) compared with 30 mM fructose (\circ) shortening. Stimulus timing is shown as a bar in b and d.

Discussion

The results extend previous data suggesting the presence of a GLUT5-like transporter, selective for fructose in guinea-pig OHCs (21). Immunohistochemical evidence shows that a GLUT5 epitope is expressed in the OHC lateral membrane (28). The developmental pattern of expression runs parallel to the development of “motor” particles in the membrane (28). The data here show that OHCs have the capacity to take up fructose selectively over glucose. This uptake is not sensitive to cytochalasin B. Although maximal fructose- and glucose-induced strains are comparable, analysis of the transport kinetics shows that fructose is transported preferentially, with a half-saturation of about 16 mM. This falls within the range of values reported in other models: from 6 mM in human enterocytes (22) to 18 mM in rat enterocytes (29).

Little is known about sugar content in the perilymphatic space. Sugars are transferred into perilymph through a blood-tissue barrier by facilitative transporters within the cochlea. Fructose

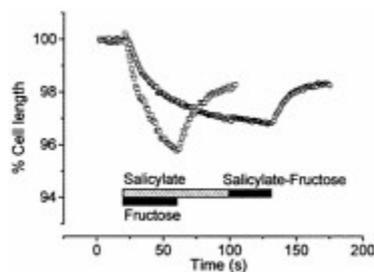


Fig. 3. Salicylate blocks fructose uptake. Pretreatment of an OHC with 10 mM sodium salicylate induced noticeable cell swelling (\circ , black) with a stabilization after 80 s. Application of a fructose-salicylate solution on such a pretreated cell failed to induce any response. After washout, the same cell treated with fructose demonstrated a fast and large change in length (\square , red). A cell length of 62 μm at the beginning of the experiment is taken as the reference for both recordings.

transport will be reduced, as blood itself does not contain more than 0.1 mM fructose, but perilymph contains about 2–4 mM glucose (30). Under these conditions, a molecule with the properties of GLUT5 cannot be using its normal substrate and it is tempting to suggest that GLUT5 is playing a different role in these cells. The most appealing hypothesis is that OHCs employ a protein in the lateral membrane in a role that is not functionally related to the protein's role in other tissues. Thus electromotility arises from (i) the high packing density of the protein and (ii) a generic property of the protein that allows it to generate areal forces in response to changes in transmembrane potential. As pointed out elsewhere (21), the rate of sugar uptake into OHCs is fast compared with other systems where it has been measured, usually with radioactive tracers. If sugar transport uptake rate is about 100 s^{-1} per carrier site, the transporter would have to be present in copy numbers exceeding 10^7 per cell to account for the rates of cell swelling. This number is comparable to the inferred number of motors in the OHC (17).

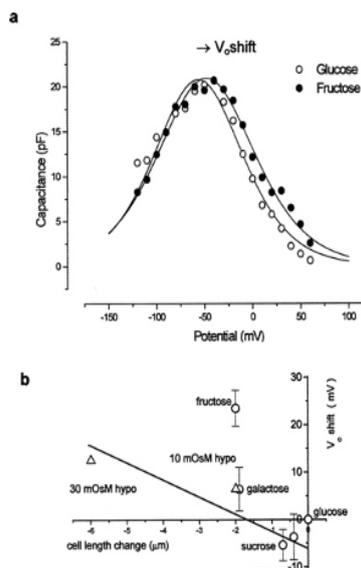


Fig. 4. OHC charge movement is modified by sugars, (a) Membrane capacitance determined by using a staircase ramp protocol. The maximum capacitance C_{\max} occurs at V_o . In this fit $\beta=0.036 \text{ mV}^{-1}$. (b) Collected data of V_o shift vs. length change for several sugars (\circ). Cell capacitance was initially recorded in 330 mosmol $\cdot\text{kg}^{-1}$ glucose. Hypoosmotic solution (Δ) also produced a V_o shift and length change. The effect of fructose shown in this plot was significantly different from that of other treatments.

As a known blocker of the OHC electromotility and charge movement (24), salicylate also affects sugar transport in guineapig OHCs. It is not clear how this molecule interacts with membrane proteins, although curiously the shape of salicylic acid bears similarities to a hexose molecule. Once within the membrane, salicylate will affect intramembranous charged particles, as salicylate induces an increase in negative surface charge (31) and behaves as a lipid-soluble anion at neutral pH. Indeed, the swelling of the cell observed when salicylate was applied suggests that it readily penetrates the cells, but at a rate lower than that of sugar. It is probable that salicylate also affects the internal membranes of the OHC, as complete reversal of the shortening was difficult after salicylate exposure.

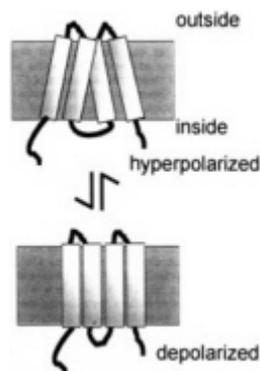


Fig. 5. Hypothetical model for a motor structure in the OHC membrane. Only four of the transmembrane helices are shown. The net area change in the plane of the membrane, produced by helix tilt, can be sufficient to produce the 5% length change observed in an OHC when electrically stimulated.

The voltage dependence of nonlinear membrane capacitance is also dependent on the ionic strength of the external solution. In whole-cell recording conditions (32), reducing the external sodium concentration shifts the position of the peak V_o of the nonlinear capacitance reversibly by up to 30 mV in the negative direction, with a mean slope of 23 mV/decade. The effect appears to be specific for sodium, because simple replacement of sodium by *N*-methyl-D-glucamine showed no effect, and replacement of C1 by gluconate was without effects on the charge movement (J.E.Gale and J.F.A., data not shown). We therefore deduce that the voltage dependence of the nonlinear capacitance depends on external ionic strength. The most economical hypothesis to explain the data is that there is a negative surface charge on the OHC. To calculate the size of the charge requires application of the Grahame equation (33). We find that including the effect of external divalent ions as well as monovalent ions requires that the OHC has a surface charge of $1e^-$ per 16 nm^2 , the negative charge accounting for the leftward shift as ionic strength is reduced.

To relate the functional studies to the structure of protein candidates is uncertain, not least because the structures of very few integral membranes are known. Like other hexose transporters, GLUT5 is a protein containing 501 amino acids in the isoform found in mouse and is significantly hydrophobic. As an integral membrane protein it is predicted to contain 12 transmembrane α -helix regions. Analysis of the primary sequence suggests that there is a charge dipole associated with the second transmembrane helix from the N terminus of the protein. What is not known is whether such hydrophobic proteins also possess

an intrinsic dipole moment. This dipole moment arises as a result of the charge associated with the peptide bonds of any α -helix (34). The dipole moment per helix is estimated to be equivalent to about $0.5 e^-$ at either end and has been suggested to contribute to the packing and stability of some proteins in membranes (35). Both of these intrinsic dipoles of GLUT5 may form part of the voltage sensor structure. In a packed assembly of molecules the total dipole would be the result of vectorial addition of individual dipole moments from all of the transmembrane portions of the helix.

Uncancelled electric charge on the protein may be the origin of the surface charge estimate of $1 e^-$ per 16 nm^2 . Since each α -helix is estimated to occupy an area of approximately 1.42 nm^2 , this surface charge would correspond to an assignment of $1 e^-$ per molecule of GLUT5, with the same estimate for any comparably sized transporter. For the moment, such intrinsic dipole sensors have not been observed directly, but calculations suggest that a generic property of high-density protein arrays in membranes is that a cell can change surface area (36). The high insertion density of particles in the OHC membrane thus confers electromotility upon the cell. Fig. 5 shows how a hypothetical rearrangement a subset of helices, by allowing them to twist against each other in the membrane, might produce the fast area change required of the OHC motor.

By subtractive PCR hybridization from OHCs and functional expression in a kidney cell line, a molecule (named "prestin") has also been identified as a candidate for the OHC motor (37). It has properties that mirror many of the anticipated properties of the hypothesized OHC motor. Isolated from a gerbil cochlear library, the protein contains 744 amino acids. This molecule, like GLUT5, does not have a characterized topology. The region identified with the charge sensor in prestin lies in a region of the molecule that, by hydrophobicity mapping alone, is in a relatively hydrophilic section of the molecule. Prestin shows homology to a family of proteins expressing a sulfate transporter motif, and thus might technically be described as an anion transporter. Physiological studies of related family members, for example the products of the *pendrin* gene, show that sulfate transport may be questionable (38) and that the family may be better described as chloride-iodide transporter (39). This molecule is a presumptive anion transporter. Nevertheless there is no evidence for sulfate transport in OHCs (J.F.A., unpublished data), and lowering extracellular chloride also has no effect on the motor (18). Neutral chloride transport can occur in OHCs, where it has been associated with bicarbonate transport (40), but at present it is unknown whether members of the "sulfate transporter" family can also use other substrates, such as a sugar.

Both of the candidate OHC motor molecules are too small by themselves to be the 8-nm-diameter motor particle observed in the freeze-etch replicas of the lateral membrane. On the basis of the size of the particles, and allowing for tissue preparation artifacts, the observed particle type is thus likely to represent a cluster of membrane-inserted proteins with a combined molecular mass of 200–250 kDa. This would match either a tetramer of GLUT5 (56 kDa) or a trimer of prestin (81 kDa). Alternatively, both molecules may be required to form a stable multiunit heteromer. In support of the latter possibility, freeze-fracture replicas of the P-face of OHCs show structures that indicate a multimeric unit (16). Although the structural topology of the motor is unclear, how the OHC inserts such high densities of protein into the basolateral membrane remains another issue that deserves attention. The insertion, assembly, and regulation of this array of molecular motors make OHCs able to carry out their role as force generators at high frequencies within the cochlear partition. How to span the gap between protein sequence and a molecular description of function must be one of the postgenomic challenges for the future.

We thank Jonathan Gale for discussion. This work was supported by the Wellcome Trust and the Medical Research Council.

1. Manley, G.A. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11736–11743.
2. Gold, T. (1948) *Proc. R. Soc. London Ser. B* **135**, 492–498.
3. Flock, A. & Cheung, H. (1977) *J. Cell Biol* **75**, 339–343.
4. Hudspeth, A.J., Choe, Y., Mehta, A.D. & Martin, P. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11765–11772.
5. Zweig, G. (1991) *J. Acoust. Soc. Am.* **89**, 1229–1254.
6. Brownell, W.E., Bader, C.R., Bertrand, D. & De Ribaupierre, Y. (1985) *Science* **227**, 194–196.
7. Ashmore, J.F. (1987) *J. Physiol.* **388**, 323–347.
8. Dallos, P. & Evans, B.N. (1995) *Science* **267**, 2006–2009.
9. Frank, G., Hemmert, W. & Gummer, A.W. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4420–4425.
10. Neely, S.T. & Kim, D.O. (1986) *J. Acoust. Soc. Am.* **79**, 1472–1480.
11. Mammano, F. & Nobili, R. (1993) *J. Acoust. Soc. Am.* **93**, 3320–3332.
12. Hemmert, W., Zenner, H.P. & Gummer, A.W. (2000) *Biophys. J.* **78**, 2285–2297.
13. Dallos, P. & Evans, B.N. (1995) *Science* **267**, 2006–2009.
14. Holley, M.C. & Ashmore, J.F. (1988) *Proc. R. Soc. London Ser. B* **232**, 413–429.
15. Kalinec, F. & Kachar, B. (1992) *Biophys. J.* **62**, A166.
16. Forge, A. (1991) *Cell Tiss. Res.* **265**, 473–483.
17. Ashmore, J.F. (1992) in *Sensory Transduction*, eds Corey, D.P. & Roper, S.D. (Rockefeller Univ. Press, New York), pp. 395–412.
18. Gale, J.E. & Ashmore, J.F. (1997) *Nature (London)* **389**, 63–66.
19. Santos-Sacchi, J. (1991) *J. Neurosci.* **11**, 3096–3110.
20. Nakazawa, K., Spicer, S.S. & Schulte, B.A. (1995) *Hear. Res.* **82**, 93–99.
21. Geleoc, G.S., Casalotti, S.O., Forge, A. & Ashmore, J.F. (1999) *Nat. Neurosci.* **2**, 713–719.
22. Burant, C.F., Takeda, J., Brot-Laroche, E., Bell, G.I. & Davidson, N.O. (1992) *J. Biol. Chem.* **267**, 14523–14526.
23. Concha, I.I., Velasquez, F.V., Martinez, J.M., Angulo, C., Droppelmann, A., Reyes, A.M., Slebe, J.C., Vera, J.C. & Golde, D.W. (1997) *Blood* **89**, 4190–4195.
24. Tunstall, M.J., Gale, J.E. & Ashmore, J.F. (1995) *J. Physiol.* **485**, 739–752.
25. Santos-Sacchi, J. (1992) *J. Neurosci.* **12**, 1906–1916.
26. Iwasa, K.H. (1993) *Biophys. J.* **64**, 102.
27. Gale, J.E. & Ashmore, J.F. (1993) *J. Physiol.* **467**, 274P.
28. Nakazawa, K., Spicer, S.S. & Schulte, B.A. (1995) *Hear. Res.* **82**, 93–99.
29. Corpe, C.P., Boveland, F.J., Hoekstra, J.H. & Burant, C.F. (1998) *Biochim. Biophys. Acta* **1402**, 229–238.
30. Dancer, A. (1992) *Audiology* **31**, 301–312.
31. Spedding, M. (1984) *Br. J. Pharmacol.* **83**, 211–220.
32. Gale, J.E. & Ashmore, J.F. (1997) *Nature (London)* **389**, 63–66.
33. Hille, B., Woodhull, A.M. & Shapiro, B.I. (1975) *Philos. Trans. R. Soc. London B* **270**, 301–318.
34. Hol, W.G. (1985) *Prog. Biophys. Mol. Biol.* **45**, 149–195.
35. Ben-Tal, N. & Honig, B. (1996) *Biophys. J.* **71**, 3046–3050.
36. Raphael, R.M., Popel, A.S. & Brownell, W.E. (2000) *Biophys. J.* **78**, 2844–2862.
37. Zheng, J., Shen, W., He, D.Z., Long, K.B., Madison, L.D. & Dallos, P. (2000) *Nature (London)* **405**, 149–155.
38. Bogazzi, F., Bartalena, L., Raggi, F., Ultimieri, F. & Martino, E. (2000) *J. Endocrinol. Invest.* **23**, 170–172.
39. Scott, D.A. & Karniski, L.P. (2000) *Am. J. Physiol.* **278**, C207–C211.
40. Ikeda, K., Saito, Y., Nishiyama, A. & Takasaka, T. (1992) *J. Physiol.* **447**, 627–648.

PUTTING ION CHANNELS TO WORK: MECHANOELECTRICAL TRANSDUCTION, ADAPTATION, AND AMPLIFICATION BY HAIR CELLS

A.J.Hudspeth*, Y.Cho, A.D.Mehta, and P.Martin

Howard Hughes Medical Institute and Laboratory of Sensory Neuroscience, The Rockefeller University, 1230 York Avenue, New York, NY 10021-6399

As in other excitable cells, the ion channels of sensory receptors produce electrical signals that constitute the cellular response to stimulation. In photoreceptors, olfactory neurons, and some gustatory receptors, these channels essentially report the results of antecedent events in a cascade of chemical reactions. The mechano-electrical transduction channels of hair cells, by contrast, are coupled directly to the stimulus. As a consequence, the mechanical properties of these channels shape our hearing process from the outset of transduction. Channel gating introduces nonlinearities prominent enough to be measured and even heard. Channels provide a feedback signal that controls the transducer's adaptation to large stimuli. Finally, transduction channels participate in an amplificatory process that sensitizes and sharpens hearing.

It takes work to open an ion channel. A voltage-sensitive cation channel associated with the action potential, for example, opens after the performance of a certain amount of electrical work in the form of gating-charge movement through the transmembrane electrical field. The gating of a ligand-activated channel, such as the receptor for a neurotransmitter, is mediated by the chemical work done on ligand binding. Finally, a mechanically activated channel, such as the transduction channel of a hair cell, opens or closes in response to mechanical work done by sound or acceleration.

Statistical thermodynamics describes the requirement for work in channel gating. If the internal-energy content of a two-state channel is E_O in the open state and E_C in the closed state, the equilibrium probabilities of these two configurations, respectively p_o and p_C , are related by the Boltzmann equation

$$\frac{p_O}{p_C} = e^{-(E_O - E_C)/(kT)} = e^{-\Delta E/(kT)},$$

[1]

in which k is the Boltzmann constant and T the temperature. ΔE is the channel's change in energy on opening—the work done in gating the channel. This equation may be rearranged to yield the channel's open probability at equilibrium,

$$p_O = \frac{1}{1 + e^{\Delta E/(kT)}},$$

[2]

which explicitly relates channel gating to the associated energy difference.

Gating of Mechanically Sensitive Channels

In hair cells of the inner ear and lateral-line system, mechanical stimulation supplies the work necessary to open transduction channels. When sound sets the cochlea's basilar membrane into resonant motion, the tectorial membrane communicates force to the mechanoreceptive hair bundles. In a semicircular canal or lateral-line organ, angular acceleration or water flow displaces a cupula that transmits force to the bundles. Finally, in the otolithic organs, the utricle and saccule, linear acceleration acts on otolithic membranes that convey force to the bundles.

In response to external force, a hair bundle pivots at its base. This deflection in turn affects the tension in elastic elements, termed gating springs (1), that affect the open probability of the mechano-electrical transduction channels (reviewed in refs. 2-4). Each gating spring is probably a tip link, a filament interconnecting adjacent stereocilia (5, 6), or the cytoskeleton associated with its insertions (3). The channel's open probability is

$$p_O = \frac{1}{1 + e^{-z(X - X_0)/(kT)}},$$

[3]

in which X represents the hair bundle's displacement from its resting position, and X_0 is the displacement at which the open probability is one-half. z is the single-channel gating force, a measure of mechanical sensitivity. If the gating spring is Hookean, $z = ykd$, in which K is the spring's stiffness and d is the distance by which the spring shortens when the channel opens. The term y denotes the geometric projection factor relating gating-spring extension to hair-bundle displacement; moving the bundle's top by a distance X extends the tip link by an amount yX (7).

A hair bundle's mechanical properties may be understood from analysis of the balance of forces within the bundle. In a resting hair bundle (Fig. 1A), contiguous stereocilia are pulled toward one another by the intervening tip link, which therefore bears tension at rest (6, 8). The actin-filled rootlet of each stereocilium, however, acts as a flexional spring that resists stereociliary deflection (9, 10). At equilibrium, the hair bundle is strained internally like a strung bow: the tension in the tip link produces a torque on the stereocilium equal and opposite that exerted by the stereociliary pivots. If the N transduction channels are arranged in parallel and if we disregard a relatively minor term reflecting channel opening at rest,

$$Nykc_C \approx K_{sp}X_s. \quad [4]$$

Here x_c is the extension of the gating spring with the channel closed, K_{sp} the combined stiffness of the stereociliary pivots, and X_s the distance that the bundle moves when the tip links are disconnected (6, 7, 11).

It is possible to estimate the values of the foregoing parameters experimentally. For a large hair bundle from the bullfrog's sacculus, electron-microscopic examination indicates that $y \approx 0.14$ (7, 12). Both ultrastructural (6, 12) and mechanical measurements (7) indicate that $N \approx 50$. When Ca^{2+}

*To whom reprint requests should be addressed. E-mail: hudspaj@rockvax.rockefeller.edu.

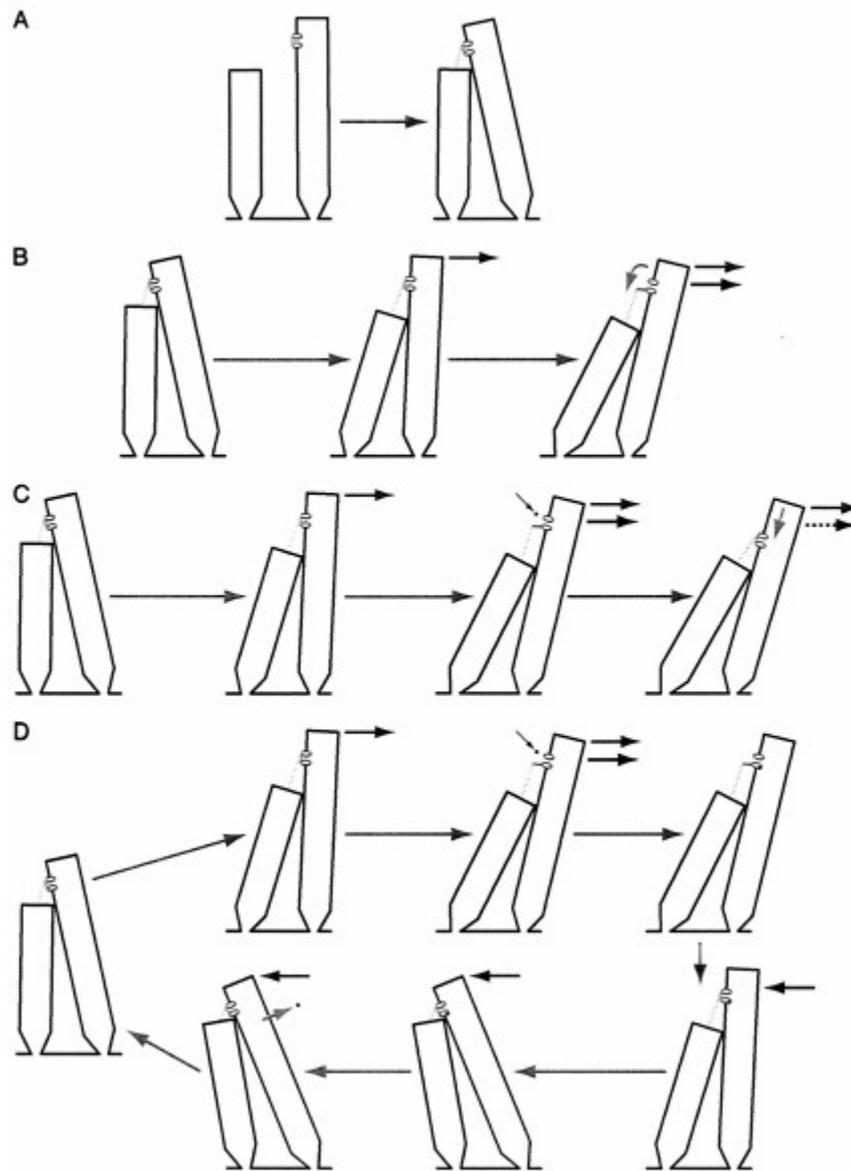


Fig. 1. Mechanical behaviors of a hair bundle. Although a bundle may contain from ≈ 20 to >300 stereocilia, it is represented here by only two. The sizes of various constituents are distorted for the sake of clarity; in reality, a stereocilium (black) is $<1 \mu\text{m}$ to $\approx 100 \mu\text{m}$ in length, a tip link (pink) is $\approx 150 \text{ nm}$ long, a channel (yellow) is $<10 \text{ nm}$ in diameter, and its gate (orange) moves by $\approx 4 \text{ nm}$. The actual movements of a hair bundle are far smaller than illustrated: a robust stimulus (60 dB sound-pressure level) deflects the bundle only $\pm 10 \text{ nm}$, half the thickness of the stereociliary outline depicted here, whereas a threshold stimulus moves the bundle less than one-tenth that far. (A) Resting strain. The stereocilia extend initially straight from the apical surface of a hair cell. When a tip link joins contiguous processes, however, it draws the stereociliary tips together and cocks the bundle in the negative direction to its resting position. The tension in the tip link then balances the strain in the actin-filled pivots at the stereociliary bases. (B) Gating compliance. Application to a hair bundle of a positively directed force (green arrow) extends the tip link. When a channel opens (curved orange arrow), the associated tip link shortens, and the tension in the link falls. Relaxation of the tip link acts like an external force in the positive direction, causing the bundle to move still further (red arrow). (C) Adaptation. A positive stimulus force (green arrow) initially deflects the hair bundle, opening a transduction channel. A Ca^{2+} ion (red) that enters through the channel interacts with a molecular motor, probably myosin $\text{I}\beta$, and causes it to slip down the stereocilium's actin cytoskeleton (orange arrow). Slackening of the tip link fosters a slow movement of the bundle in the positive direction (dashed red arrow). The reduced tension in the tip link then permits the channel to reclose. (D) Amplification. When a hair bundle is deflected by the positive phase of a sinusoidal stimulus (upper green arrows), channel opening facilitates bundle movement (upper red arrow). A Ca^{2+} ion (red) that enters through the transduction channel binds to a cytoplasmic site on or associated with the channel, promoting its reclosure. As the channel shuts, the increased tension in the tip link exerts a force that moves the bundle in the negative direction (lower red arrow), enhancing the effect of the negatively directed phase of a stimulation (lower green arrows). When the ion is extruded by a membrane Ca^{2+} pump, or Ca^{2+} -ATPase (orange arrow), the hair bundle is primed to repeat the cycle.

chelators sever tip links, a hair bundle moves in the positive direction to the equilibrium position of the stereociliary pivots; for cells originally in a medium containing 4 mM Ca^{2+} , $X_0 \approx 130$ nm (6). Stiffness measurements in the absence of tip links or orthogonal to them indicate that $K_{\text{SP}} \approx 200 \mu\text{N}\cdot\text{m}^{-1}$ (13, 14); determinations of the hair bundle's total stiffness then imply that $K \approx 700 \mu\text{N}\cdot\text{m}^{-1}$ (14). Analysis of bundle properties during channel gating indicate that $d \approx 4$ nm (7, 15; but see ref. 16). Finally, calculations from the foregoing results suggest that $z \approx 0.6$ pN and $xc \approx 7$ nm for hair bundles exposed to 4 mM Ca^{2+} .

When a hair bundle is deflected by mechanical stimulation, the shearing motion between contiguous stereocilia produces a force that in each tip link affects the gating of the attached channel (Fig. 1B). The external force, F_{HB} , required to hold the hair bundle at a particular position is given by (3, 11)

$$F_{\text{HB}} = Nyk(yX + xc - p_o d) + K_{\text{SP}}(X - X_0) = (Ny^2 K + K_{\text{SP}})X - Np_o z + F_o, \quad [5]$$

in which F_o is an offset term such that F_{HB} is zero at the bundle's resting position. This relation includes a critically important negative term that associates a decrease in the applied force with an increase in the channel open probability p_o . When a transduction channel opens, the associated tip link relaxes by the distance d (Fig. 1B). The force exerted by the link, which pulls the hair bundle in the negative direction at rest, therefore declines on opening by d —equivalent to exerting a positive force on the bundle. Measured at the bundle's top, the ensemble of channels can produce a force as great as $N\gamma d$. The system thus displays positive feedback: a stimulus that moves a hair bundle in the positive direction opens transduction channels, whose opening in turn promotes additional positive movement. As we shall see below, nature has evidently exploited this characteristic to do useful work.

The hair bundle's stiffness may be calculated from the foregoing expression to be

$$K_{\text{HB}} = \frac{dF_{\text{HB}}}{dX} = N\gamma^2 \kappa + K_{\text{SP}} - \left(\frac{Nz^2}{kT} \right) p_o (1 - p_o), \quad [6]$$

in which the open probability depends on bundle displacement as described by Eq. 3 (7, 11). The stiffness (Fig. 2A) is nearly constant for displacements well in the negative direction, for which $p_o \approx 0$, or positive direction, for which $p_o \approx 1$. Each increment in the force exerted on the hair bundle, ΔF_{HB} , accordingly produces an approximately constant change in bundle position, ΔX :

$$\Delta X \approx \left(\frac{1}{N\gamma^2 \kappa + K_{\text{SP}}} \right) \Delta F_{\text{HB}}. \quad [7]$$

Over the range of displacements in which the channels gate, though, the stiffness declines. At the position where half the channels are open, X_o , the stiffness reaches its minimal value, and the bundle's sensitivity to stimulus force peaks:

$$\Delta X \approx \left[\frac{1}{N\gamma^2 \kappa + K_{\text{SP}} - \left(\frac{Nz^2}{4kT} \right)} \right] \Delta F_{\text{HB}}. \quad [8]$$

This effect, known as gating compliance, has been observed experimentally in hair cells of the frog's sacculus (7) and probably in those of the teleost lateral-line organ (17), mammalian cochlea (18), and mammalian vestibule (19). The hair bundle's mechanical nonlinearity can account for distortion products (20), the illusory tones heard when sinusoidal stimuli of two or more frequencies are presented concurrently. These distortions seem to be an ineluctable price paid for the advantages—speed and sensitivity—associated with the direct transduction of mechanical stimuli.

Note that gating compliance could actually render the hair bundle's stiffness negative over a certain range of displacements (4, 15) if

$$\frac{Nz^2}{4kT} > N\gamma^2 \kappa + K_{\text{SP}}. \quad [9]$$

Because the final term is substantially smaller than the penultimate one, the definition of z allows this relation to be restated as

$$d^2 \gtrsim 4kT. \quad [10]$$

If the stiffness were locally negative, the unrestrained hair bundle would be bistable (4), residing at either of two stable points flanking the negative-slope region. Under these circumstances, the bundle could nevertheless transduce small stimuli if thermal noise were sufficient to surmount the energy barrier separating the two states, a phenomenon termed stochastic resonance (21). Moreover, if supplemented with an energy source to bias its operation into the unstable region, such a hair bundle would potentially be capable of producing active oscillations.

Comparison of Electrically and Mechanically Sensitive Channels

It is interesting to contrast the behavior of a mechanically stimulated ion channel with that of a channel responsive to electrical stimuli. Consider, for example, a member of the well-studied superfamily of voltage-activated Na^+ , Ca^{2+} , and K^+ channels. When the charge on the interior surface of the cell's membrane becomes more positive, such a channel opens as one or more gating moieties undergo a rearrangement equivalent to the movement of ζ charges through a transmembrane potential V_M . The open probability in this instance is

$$p_o = \frac{1}{1 + e^{-\zeta e(V_M - V_o)/(kT)}}, \quad [11]$$

in which e is the elementary charge and V_o is the potential at which the open probability is one-half (reviewed in ref. 22).

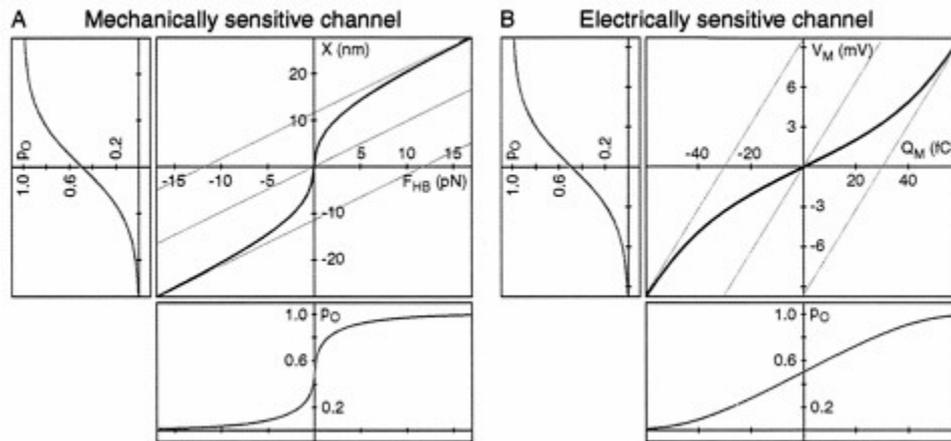


Fig. 2. Gating of ion channels. (A) Gating of a mechanically sensitive ion channel. The large plot relates hair-bundle displacement (X) to the external force applied to the bundle (F_{HB}). As a result of gating compliance, channel gating is highly sensitive over a narrow range of forces. The green lines portray the linear relations expected if the channels were to remain closed (top line), open (bottom line), or both in equal numbers (middle line). The curve was obtained by numerically solving a transcendental equation, the inverse of Eq. 5; the parameter values were $n=35$, $\gamma=0.14$, $\kappa=1200 \mu\text{N}\cdot\text{m}^{-1}$, $d=4 \text{ nm}$, $z=0.67 \text{ pN}$, and $K_{SP}=200 \mu\text{N}\cdot\text{m}^{-1}$. The plot at the left represents the Boltzmann relation between displacement and the channel's open probability (Eq. 3); the displacement axis is identical to that in the principal graph. The graph at the bottom displays the dependence of the channel's open probability on hair-bundle force; the force axis accords with that in the main plot. Note the steepness of this relation, in comparison both to that of a Boltzmann curve and especially to that of the corresponding plot for an electrically activated channel in B. (B) Gating of an electrically sensitive ion channel. The large graph relates membrane potential (V_M) to the charge applied to the membrane (Q_M). As a consequence of gating capacitance, the membrane potential is less responsive to charge over the range of potentials in which channels open and close. The green lines portray the linear relations expected if the channels were to remain closed (left line), open (right line), or both in equal numbers (center line). The curve was obtained by numerically solving a transcendental equation, the inverse of Eq. 12; the numerical values were those provided in the text, with $Q_0=-30 \text{ fC}$. The plot at the left represents the Boltzmann relation between membrane potential and the channel's open probability (Eq. 11); the voltage axis is identical to that in the principal graph. The graph at the bottom displays the dependence of the channel's open probability on the charge applied to the membrane; the charge axis accords with that in the main plot. To facilitate comparison, the four open-probability plots in A and B were scaled such that each extends from $p_o=0.01$ to $p_o=0.99$.

As a voltage-activated channel progresses from one configuration to another, say from the closed to the open state, it exchanges a certain amount of energy with the environment in the form of charge displacement across the membrane. For any membrane potential, the total membrane charge, Q_M , comprises that deposited on the lipid bilayer, Q_B , and the gating charge, Q_G , associated with the opening and closing of an ensemble of N channels:

$$\begin{aligned} Q_M &= Q_B + Q_G + Q_0 \\ &= C_B V_M + N p_o z e + Q_0 \\ &= C_B V_M + \frac{N z e}{1 + e^{-z e (V_M - V_0) / (kT)}} + Q_0. \end{aligned} \quad [12]$$

Here C_B represents the capacitance caused by the lipid bilayer, whose value per unit area is typically $\approx 10 \text{ mF}\cdot\text{m}^{-2}$. Q_0 is an offset term such that Q_M is zero when the membrane is electrically unpolarized. It follows that the system's capacitance, C_M , is not constant (23), for

$$C_M = \frac{dQ_M}{dV_M} = C_B + \left(\frac{N z^2 e^2}{kT} \right) p_o (1 - p_o). \quad [13]$$

The final term represents gating capacitance, the increase in the membrane's capacitance because of the movement of gating moieties, whose magnitude depends on the open probability and hence on the membrane potential.

The impact of the gating charge on neural signaling may be appreciated by examining the relation between the charge applied to the membrane and the ensuing membrane potential (Fig. 2B). For extreme hyperpolarization, during which $p_o \approx 0$, or depolarization, during which $p_o \approx 1$, each increment of charge, ΔQ_M , produces an essentially constant voltage change, ΔV_M :

$$\Delta V_M \approx \left(\frac{1}{C_B} \right) \Delta Q_M. \quad [14]$$

Over the range of potentials at which the channels gate, however, the membrane capacitance increases. The voltage therefore changes less per unit of charge deposited on the membrane. At the potential V_o , the midpoint of the gating range, the responsiveness reaches its minimal value,

$$\Delta V_M \approx \left[\frac{1}{C_B + \left(\frac{N z^2 e^2}{4kT} \right)} \right] \Delta Q_M. \quad [15]$$

The capacitive load imposed by gating can be substantial. Consider an isopotential, spherical neuron $10 \mu\text{m}$ in diameter. Suppose that this cell is endowed at the modest density of $\approx 40 \mu\text{m}^{-2}$ with identical voltage-sensitive channels, each with an effective gating charge of $12e$ (24). The maximal capacitance

contributed by channel gating then equals that caused by the membrane bilayer! After an excitable cell has reached threshold for firing of an action potential, the capacitive cost imposed by channel gating is repaid by the far larger current supplied through ion channels.

Gating capacitance also has considerable impact on the speed of neural signaling. An organism is likely to enjoy a selective advantage if its nervous system operates with particular rapidity. Speed is important in neural computation, which consists of integrating synaptic inputs until the threshold is reached for initiating an action potential. In addition, the propagation velocity of the action potential determines how quickly information may be moved between cells in the nervous system. When ionic current flows through activated synaptic receptors, gating capacitance slows synaptic potentials and thus the approach to threshold. Moreover, after an action potential has been initiated in an axon, gating charge diminishes the rate at which longitudinal current can depolarize successive increments of the axonal membrane. This effect on propagation is great enough that an excessive density of channels in the membrane of an unmyelinated fiber would actually retard the signal (23, 25).

The energetic cost of gating is shared by electrically and mechanically activated channels (Fig. 2). The symmetry in the descriptions of the two gating mechanisms is broken, however, by the nature of the relevant independent parameters, charge and force, and the dependent parameters, voltage and displacement. This distinction introduces a critical difference in the signs of the final terms in the denominators of Eqs. 8 and 15. The gating of voltage-activated channels *increases* membrane capacitance and therefore constitutes negative feedback that acts to retard change in the membrane potential. The gating of force-activated channels, by contrast, *decreases* hair-bundle stiffness and thus provides positive feedback that promotes bundle motion. The distinct behaviors of the two types of channel are apparent from the dependence of channel open probability on the respective inputs, electrical charge and mechanical force (Fig. 2). By broadening the relation between charge and open probability, gating capacitance desensitizes electrically activated cells. Gating compliance, on the other hand, narrows the relation between force and open probability and thereby enhances sensitivity in mechanically responsive cells.

Gating compliance is an intrinsic property of direct mechano-electrical transduction that natural selection may well have heightened to increase the ear's sensitivity. A hair bundle's responsiveness to force improves as the denominator of Eq. 8 approaches zero. Sensitivity could be enhanced by increasing the single-channel gating force, z , either by augmenting γ , the stiffness of the gating spring, or by increasing d , the distance by which the spring shortens on channel opening. From the evolutionary perspective, the former change might be effected by mutations that produce a gating-spring protein that is more rigid in extension. If d represents the "swing" of the channel's gate, the moiety that obstructs the ion-conducting pore, its magnitude would be constrained by the size of the channel molecule. If the gate were to evolve a lever-like extension, however, the value of d might increase without disrupting channel structure (3).

Adaptation

The hair cell's great sensitivity poses a problem: how might the mechanoreceptor's response be prevented from saturating in the presence of large static stimuli? In other words, if mechano-electrical transduction is direct, without the intervention of a second messenger between the detection of force and the opening of channels, how can the transducer's sensitivity be adjusted? The answer to these questions involves a novel form of sensory adaptation (reviewed in refs. 26–29).

When a hair bundle from the frog's sacculus is deflected in the positive direction, an initial surge of inward transduction current rapidly depolarizes the hair cell. Within milliseconds, however, the response declines, following an approximately exponential time course to a plateau (30–32). Negative stimulation elicits similar adaptation; an immediate decrease in the transduction current and the associated hyperpolarization are followed by a gradual return toward the resting state. The association of adaptation with mechanical relaxation of the hair bundle (7, 8, 13, 33) initially hinted that the process is mediated by a physical readjustment of tension in the gating springs. It was accordingly hypothesized that the position of each tip link's upper insertion is adjusted during adaptation (13, 33): the link's insertion slides downward in response to positive stimulation (Fig. 1C) and climbs upward after negative stimuli. In view of the actin cytoskeleton of stereocilia, a protein of the myosin family was proposed as the motor to drive adaptation.

Several lines of evidence now buttress this model for adaptation. Along with at least three other myosin isoforms, hair bundles contain myosin I β (34, 35), which is concentrated near the ends of each tip link (36, 37). The number of myosin molecules present in a stereocilium (34) could easily account for the forces produced by the adaptation motor (8). Whole-cell dialysis of hair cells indicates that the adaptation motor is inhibited by ATP depletion (38) and by phosphate analogs (39), both characteristics of myosin. Adaptation is regulated by Ca²⁺ (31, 32), which readily traverses the transduction channel (40, 41), accumulates in the stereociliary cytoplasm (42) and is extruded by Ca²⁺ pumps (43). The Ca²⁺ sensor for adaptation appears to be calmodulin (44), which occurs at a high concentration near stereociliary tips (45, 46) and can occupy the light-chain binding sites on myosin I β (47, 48). Finally, the biochemical properties of myosins I (reviewed in ref. 49) accord with the activity of the adaptation motor. During steady-state ATP turnover, these mechanoenzymes characteristically detach from actin more rapidly in the presence of Ca²⁺ (50), which might in the present context explain the downward slippage of the adaptation motor when a channel is open.

The effects of adaptation on channel open probability and on the hair bundle's mechanical properties can be accommodated readily by the gating-spring model (32, 51; reviewed in ref. 3). When a tip link's upper insertion moves down the stereocilium by a distance x_A , the open probability of the associated channel falls to

$$p_o = \frac{1}{1 + e^{-z(X - X_0 - x_A/\gamma)/(kT)}} \quad [16]$$

In other words, adaptation offsets the setpoint of transduction by an amount x_A/γ , as measured atop the bundle. The force necessary to hold the bundle at displacement X correspondingly declines to

$$F_{HB} = Nyk(yX + X_C - X_A - p_o d) + K_{SP}(X - X_s) \quad [17]$$

Adaptation does not proceed to completion, though; for stimuli of small to moderate size, mechanosensitivity generally migrates $\approx 80\%$ of the magnitude of an imposed bundle offset (30, 31, 51).

Adaptation serves at least four ends. Every hair cell may need the process to avoid saturation of its responsiveness by large stimuli. Next, despite irregularities in hair-bundle growth and in the operation of the conductive apparatus that delivers stimuli to a hair cell, adaptation situates a bundle in a sensitive region of its operating range. Because cells with especially rapid adaptation are insensitive to low-frequency stimuli, adaptation participates in high-pass filtering during the processing of sensory inputs (52). As a consequence of dissimilar rates of adaptation, cells in a single receptor organ may differ in their frequency responsiveness (37, 53). Finally, as discussed below, adaptation

poises the transduction machinery so that the hair bundle is most capable of doing active mechanical work (54).

In keeping with the numerous roles of adaptation, the process is common to hair bundles in many vertebrates. As detailed above, adaptation is prominent in hair cells of the amphibian sacculus. Adaptation also occurs in the amphibian utricle, where cells of disparate bundle morphology display strikingly different rates (53). The corresponding process in reptiles, termed slow adaptation (55, 56), resembles adaptation in the frog but is not thought to be mediated by a mechanical readjustment of the hair bundle (57). In mammals, adaptation has been documented electrophysiologically in hair cells of the murine cochlea and utricle (28, 58, 59). In addition, the mechanical relaxation of the hair bundle associated with adaptation occurs in both organs (19, 60).

Amplification

Most of the foregoing discussion treats the hair bundle as a passive structure that is deflected by stimuli. As is apparent from the mechanical correlates of adaptation, however, a hair cell is also capable of actively producing forces. This means that a hair bundle can do work against an external load with energy provided by the hair cell itself. This capacity raises two important questions: what purpose is served by active hair-bundle work, and what cellular motor accomplishes it?

Active hair-bundle movement may underlie the phenomenon of cochlear amplification (reviewed in refs. 61–63). Despite the damping effects of fluid in the inner ear, the cochlea acts as a resonator with a high quality factor; this behavior implies the presence in the ear of an active process that counters viscous dissipation (64). A large body of research (reviewed in ref. 63) indicates that a mechanical amplifier confers on the inner ear four linked properties: enhanced mechanical sensitivity, sharpened frequency selectivity, metabolic vulnerability, and the ability to produce spontaneous otoacoustic emissions. These features are clearly present in cochleas, the hearing organs of reptiles, birds, and mammals. Moreover, the primitive auditory receptor organs of amphibians display similar characteristics; those of fishes have not been tested. The phenomenon of electromotility or voltage-driven change in the length of an outer hair cell is believed to constitute the active process in the mammalian cochlea (reviewed in refs. 65, 66, but see ref. 67). Because electromotility seems not to occur in nonmammalian tetrapods, however, there must be an additional active process in these animals. The resemblance of mammalian audition to that in other vertebrates (reviewed in ref. 63) may alternatively imply the existence of a ubiquitous active process, present in all tetrapods and perhaps supplemented by electromotility to permit the higher-frequency hearing of mammals.

In addition to the slower motions associated with adaptation (32, 68), hair bundles can evince several types of active mechanical behavior. Some bundles display spontaneous oscillations (9, 13) that are severalfold as great as expected for the thermal vibration of passive bundles of the stiffness measured (69). The narrow frequency spectra of these movements (68) suggest that they reflect the operation of a tuned mechanical oscillator. Active motions of a second type are evoked by abruptly deflecting a hair bundle with a flexible fiber. In hair cells of the turtle's cochlea, this stimulus elicits damped mechanical oscillations (9). When maintained in saline solution containing 4 mM Ca^{2+} , stimulated hair cells from the frog's sacculus produce abrupt twitches (13, 54). If the Ca^{2+} concentration is reduced to the level normally found in endolymph, however, the response becomes slower and more resonant (54, 70). Damped oscillatory responses also occur in the short hair cells of the avian cochlea (Fig. 3), which are specialized for mechanical amplification (reviewed in refs. 62, 63, 71). Rapid, active hair-bundle movements have not been reported from the inner ears of fishes or mammals.

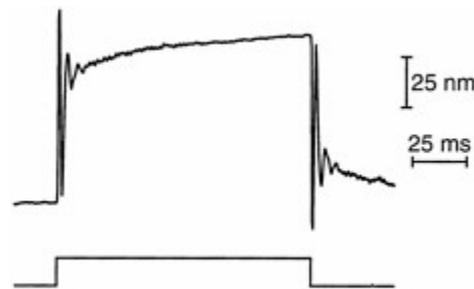


Fig. 3. Mechanically evoked response of a hair bundle from the chicken's cochlea. A damped sinusoidal oscillation of the hair bundle (*Upper*) ensues from stimulation of a short hair cell situated roughly one-quarter of the distance along the cochlea from its apex. This active mechanical response resembles those recorded earlier from hair cells of the turtle's cochlea and the frog's sacculus, but is severalfold as large and occurs at a higher frequency of ≈ 235 Hz. An epithelial preparation of the basilar papilla was maintained at room temperature in an oxygenated archosaur saline solution including 1 mM Ca^{2+} . The tip of a flexible glass fiber, with a stiffness of $680 \mu\text{N}\cdot\text{m}^{-1}$ and a response time constant of $30 \mu\text{s}$, was attached to the bundle's top; stimulation was accomplished by abruptly displacing the fiber's base by 400 nm for 120 ms (*Lower*). The response represents the output of a dual photodiode onto which an image of the fiber's tip was projected. ($\times 1,000$.)

Perhaps the most interesting responses occur in anuran hair cells kept in a two-compartment experimental chamber, so that their apical surfaces may be bathed in artificial endolymph while their basolateral membranes are exposed to a perilymph-like medium. Under these conditions, which resemble those found in the intact ear, many bundles display spontaneous oscillatory motions at frequencies of 5–40 Hz and with amplitudes up to 40 nm or occasionally even more (72). When a sinusoidal force is exerted on such a bundle by a flexible fiber, the stimulus can entrain the bundle's motion. Entrainment is most effective at the frequency of spontaneous oscillation; for progressively lower or higher frequencies, stimuli of increasing magnitude are necessary.

Power gain, an increase in the energy content of an output signal with respect to the otherwise similar input signal, provides a benchmark for amplification. Analysis of the work done by hair bundles from the frog's ear confirms that these organelles accomplish active mechanical amplification (72). In a healthy cell, application of a sinusoidal stimulus force only a few piconewtons in amplitude can entrain the bundle's motion. During each cycle of motion, viscous drag on the bundle and the attached stimulus fiber irreversibly dissipates an amount of energy that can be calculated from the speed of bundle motion and the system's drag coefficient. Remarkably, the average viscous dissipation often exceeds the energy supplied by the stimulus fiber. The balance of energy must come from the hair bundle, which therefore does mechanical work in such a way that it amplifies the weak stimulus.

The second important question about hair bundle-based mechanical amplification is the identity of the motor that mediates the process. Two candidates have been considered. First, the myosin-based adaptation motor might also power the active process (reviewed in ref. 26). Especially if the myosin heads do not detach from actin filaments but rock in place like those of insect flight muscle, the adaptation motor might produce bundle oscillations at frequencies into the kilohertz range. Even faster movements could arise through the interplay of the collective

activity of a motor ensemble with the hair bundle's passive mechanical properties (73). The principal difficulty with this mechanism is its seeming variance with correlated mechanical and electrophysiological observations on the bullfrog's saccular hair cell (14, 54). In this preparation, positive stimulation with a flexible fiber causes hair-bundle motion in the same direction that is interrupted by a transient movement, or twitch, in the opposite direction. The initial component of motion is accompanied by inward transduction current and cellular depolarization; the twitch is associated with abrupt channel reclosure and repolarization toward the resting potential. If the adaptation motor is in series with the gating spring (reviewed in refs. 26–29), the motor must step up the stereocilium to produce a negatively directed twitch. Because such a motion increases tension in the tip link, however, it would be expected to open rather than to close transduction channels. Despite this objection, a modest relaxation of tension in the gating spring, caused for example by the adaptation motor's slippage down the stereocilium, may suffice to trigger a larger and oppositely directed movement associated with channel reclosure.

The hair-bundle twitch and related active motions may be explained by an alternative mechanism of force production in the hair bundle, fast channel reclosure owing to Ca^{2+} binding. For hair cells of the frog and turtle, the transduction channel's open probability declines with increasing Ca^{2+} concentration in the solution bathing the hair bundle (31, 52, 55). Moreover, the hair-bundle twitch accelerates as the Ca^{2+} level climbs (54, 70). These observations led to the suggestion (7, 70) that Ca^{2+} entering through an open transduction channel can combine with a site on or associated with the channel, causing it to reclose (Fig. 1D). Closure then increases the tension in the tip link, jerking the hair bundle in the negative direction. Developed to describe responses in the frog's sacculus, this model also explains channel reclosure—termed fast adaptation—in the turtle's cochlea (56, 57). In light of the similarities between the two preparations, it seems probable that measurement of the mechanical signals associated with fast channel reclosure in the turtle would yield results similar to those analyzed in the frog.

Ca^{2+} -mediated channel reclosure could power active hair-bundle motions, including both spontaneous oscillations and those entrained by a stimulus force (reviewed in ref. 61). In this model (Fig. 1D), the positive phase of a sinusoidal stimulus force initiates the opening of transduction channels. The gating compliance associated with channel opening provides a positively directed force that augments the movement. As Ca^{2+} flows into the stereociliary cytoplasm, however, it promotes channel reclosure, tightens tip links, and thus causes a negatively directed force on the hair bundle. If this response occurs with an appropriate latency, it accentuates the effect of the negative phase of stimulation.

Implemented as a computational model, the foregoing scheme can account for several properties of the cochlear amplifier (74). The proposed mechanism provides an amplification of about 100-fold by comparison with the hair bundle's passive responsiveness. The amplification is strongly level dependent: threshold stimulation evokes the greatest gain, then sensitivity falls as the stimulus grows. Finally, the amplification is sharply frequency selective. These three properties accord with the known characteristics of the cochlear amplifier (reviewed in refs. 65, 75). The model can additionally be made to produce spontaneous oscillations, which might underlie the spontaneous otoacoustic emissions that are widely associated with the inner ear's active process (reviewed in refs. 63, 76). This ability to function as a sharply tuned high-gain amplifier or as a spontaneous oscillator is intrinsic to a broad class of models, such as the present one, that display a Hopf bifurcation (73, 77).

Conclusion

The work required to open and close ion channels is an inevitable price of signaling by excitable cells. Uniquely among known sensory receptors, hair cells exploit the nonlinearity inherent in channel gating to sharpen their responsiveness to stimulation. By coupling channel gating to a power source, some hair cells have additionally developed an amplifier that augments the ear's mechanical sensitivity and frequency selectivity.

Y.C. is supported in part by a training grant from the National Science Foundation. A.D.M. and P.M. are Associates and A.J.H. is an Investigator of the Howard Hughes Medical Institute. The original research from our laboratory discussed in this article was supported by National Institutes of Health Grants DC00241 and DC00317.

- Corey, D.P. & Hudspeth, A.J. (1983) *J. Neurosci.* **3**, 962–976.
- Hudspeth, A.J. (1989) *Nature (London)* **341**, 397–404.
- Hudspeth, A.J. (1992) in *Sensory Transduction*, eds. Corey, D.P. & Roper, S.D. (Rockefeller Univ. Press, New York), pp. 357–370.
- Markin, V.S. & Hudspeth, A.J. (1995) *Annu. Rev. Biophys. Biomol. Struct.* **24**, 59–83.
- Pickles, J.O., Comis, S.D. & Osborne, M.P. (1984) *Hear. Res.* **15**, 103–112.
- Assad, J.A., Shepherd, G.M.G. & Corey, D.P. (1991) *Neuron* **7**, 985–994.
- Howard, J. & Hudspeth, A.J. (1988) *Neuron* **1**, 189–199.
- Jaramillo, F. & Hudspeth, A.J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1330–1334.
- Crawford, A.C. & Fettiplace, R. (1985) *J. Physiol.* **364**, 359–379.
- Howard, J. & Ashmore, J.F. (1986) *Hear. Res.* **23**, 93–104.
- Howard, J., Roberts, W.M. & Hudspeth, A.J. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 99–124.
- Jacobs, R.A. & Hudspeth, A.J. (1990) *Cold Spring Harbor Symp. Quant. Biol.* **55**, 547–561.
- Howard, J. & Hudspeth, A.J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3064–3068.
- Marquis, R.E. & Hudspeth, A.J. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 11923–11928.
- Denk, W., Keolian, R.M. & Webb, W.W. (1992) *J. Neurophysiol.* **68**, 927–932.
- Denk, W., Holt, J.R., Shepherd, G.M.G. & Corey, D.P. (1995) *Neuron* **15**, 1311–1321.
- van Netten, S.M. & Khanna, S.M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1549–1553.
- Russell, I.J., Kössl, M. & Richardson, G.P. (1992) *Proc. R. Soc. London Ser. B* **250**, 217–227.
- Géléoc, G.S.G., Lennan, G.W.T., Richardson, G.P. & Kros, C.J. (1997) *Proc. R. Soc. London Ser. B* **264**, 611–621.
- Jaramillo, F., Markin, V.S. & Hudspeth, A.J. (1993) *Nature (London)* **364**, 527–529.
- Jaramillo, F. & Wiesenfeld, K. (1998) *Nat. Neurosci.* **1**, 384–388.
- Hille, B. (1992) *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA), 2nd Ed., pp. 54–57.
- Hodgkin, A. (1975) *Philos. Trans. R. Soc. London B* **270**, 297–300.
- Schoppa, N.E., McCormack, K., Tanouye, M.A. & Sigworth, F.J. (1992) *Science* **255**, 1712–1715.
- Adrian, R.H. (1975) *Proc. R. Soc. London Ser. B* **189**, 81–86.
- Hudspeth, A.J. & Gillespie, P.G. (1994) *Neuron* **12**, 1–9.
- Gillespie, P.G. & Corey, D.P. (1997) *Neuron* **19**, 955–958.
- Eaton, R.A. (2000) *Annu. Rev. Neurosci.* **23**, 285–314.
- Holt, J.R. & Corey, D.P. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11730–11735.
- Eaton, R.A., Corey, D.P. & Hudspeth, A.J. (1987) *J. Neurosci.* **7**, 2821–2836.
- Hacohen, N., Assad, J.A., Smith, W.J. & Corey, D.P. (1989) *J. Neurosci.* **9**, 3988–3997.
- Assad, J.A. & Corey, D.P. (1991) *J. Neurosci.* **12**, 3291–3309.
- Howard, J. & Hudspeth, A.J. (1987) in *Sensory Transduction*, eds. Hudspeth, A.J., MacLeish, P.R., Margolis, F.L. & Wiesel, T.N. (Fondation pour l'Etude du Système Nerveux Central et Périphérique, Geneva), pp. 138–145.
- Gillespie, P.G., Wagner, M.C. & Hudspeth, A.J. (1993) *Neuron* **11**, 581–594.
- Hasson, T., Gillespie, P.G., Garcia, J.A., MacDonald, R.B., Zhao, Y., Yee, A.G., Mooseker, M.S. & Corey, D.P. (1997) *J. Cell Biol.* **137**, 1287–1307.
- Garcia, J.A., Yee, A.G., Gillespie, P.G. & Corey, D.P. (1998) *J. Neurosci.* **18**, 8637–8647.
- Steyger, P.S., Gillespie, P.G. & Baird, R.A. (1998) *J. Neurosci.* **18**, 4603–4615.
- Gillespie, P.G. & Hudspeth, A.J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2710–2714.
- Yamoah, E.N. & Gillespie, P.G. (1996) *Neuron* **17**, 523–533.
- Lumpkin, E.A., Marquis, R.E. & Hudspeth, A.J. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 10997–11002.

41. Ricci, A.J. & Fettiplace, R. (1998) *J. Physiol.* **506**, 159–173.
42. Lumpkin, E.A. & Hudspeth, A.J. (1998) *J. Neurosci.* **18**, 6300–6318.
43. Yamoah, E.N., Lumpkin, E.A., Dumont, R.A., Smith, P.J.S., Hudspeth, A.J. & Gillespie, P.G. (1998) *J. Neurosci.* **18**, 610–624.
44. Walker, R.G. & Hudspeth, A.J. (1996) *Proc. Natl Acad. Sci. USA* **93**, 2203–2207.
45. Shepherd, G.M.G., Barres, B.A. & Corey, D.P. (1989) *Proc. Natl Acad. Sci. USA* **86**, 4973–4977.
46. Walker, R.G., Hudspeth, A.J. & Gillespie, P.G. (1993) *Proc. Natl Acad. Sci. USA* **90**, 2807–2811.
47. Metcalf, A.B., Chelliah, Y. & Hudspeth, A.J. (1994) *Proc. Natl Acad. Sci. USA* **91**, 11821–11825.
48. Sole, C.K., Derfler, B.H., Duyk, G.M. & Corey, D.P. (1994) *Aud. Neurosci.* **1**, 63–75.
49. Barylko, B., Binns, D.D. & Albanesi, J.P. (2000) *Biochim. Biophys. Acta* **1496**, 23–35.
50. Coluccio, L.M. & Geeves, M.A. (2000) *J. Biol Chem.* **274**, 21575–21580.
51. Shepherd, G.M.G. & Corey, D.P. (1994) *J. Neurosci.* **14**, 6217–6229.
52. Corey, D.P. & Hudspeth, A.J. (1983) *J. Neurosci.* **3**, 942–961.
53. Baird, R.A. (1992) *Ann. N.Y. Acad. Sci.* **656**, 12–26.
54. Benser, M.E., Marquis, R.E. & Hudspeth, A.J. (1996) *J. Neurosci.* **16**, 5629–5643.
55. Crawford, A.C., Evans, M.G. & Fettiplace, R. (1991) *J. Physiol* **434**, 369–398.
56. Wu, Y.-C., Ricci, A.J. & Fettiplace, R. (1999) *J. Neurophysiol.* **82**, 2171–2181.
57. Crawford, A.C., Evans, M.G. & Fettiplace, R. (1989) *J. Physiol.* **419**, 405–434.
58. Kros, C.J., Rüscher, A. & Richardson, G.P. (1992) *Proc. R. Soc. London Ser. B* **249**, 185–193.
59. Holt, J.R., Corey, D.P. & Eatock, R.A. (1997) *J. Neurosci.* **17**, 8739–8748.
60. Russell, I.J., Richardson, G.P. & Kössl, M. (1989) *Hear. Res.* **43**, 55–70.
61. Hudspeth, A.J. (1997) *Curr. Opin. Neurobiol* **7**, 480–486.
62. Manley, G.A. & Köppl, C. (1998) *Curr. Opin. Neurobiol.* **8**, 468–474.
63. Manley, G. (2000) *Proc. Natl Acad. Sci. USA* **97**, 11736–11743.
64. Gold, T. (1948) *Proc. R. Soc. London Ser. B* **135**, 492–498.
65. Dallos, P. (1992) *J. Neurosci.* **12**, 4575–4585.
66. Nobili, R., Mammano, F. & Ashmore, J. (1998) *Trends Neurosci.* **21**, 159–167.
67. Yates, G.K. & Kirk, D.L. (1998) *J. Neurosci.* **18**, 1996–2003.
68. Denk, W. & Webb, W.W. (1992) *Hear. Res.* **60**, 89–102.
69. Denk, W., Webb, W.W. & Hudspeth, A.J. (1989) *Proc. Natl Acad. Sci. USA* **86**, 5371–5375.
70. Jaramillo, F., Howard, J. & Hudspeth, A.J. (1990) in *The Mechanics and Biophysics of Hearing*, eds. Dallos, P., Geisler, C.D., Matthews, J.W., Ruggero, M.A. & Steele, C.R. (Springer, Berlin), pp. 26–33.
71. Fettiplace, R. & Fuchs, P.A. (1999) *Annu. Rev. Physiol.* **61**, 809–834.
72. Martin, P. & Hudspeth, A.J. (1999) *Proc. Natl Acad. Sci. USA* **96**, 14306–14311.
73. Camalet, S., Duke, T., Jülicher, F. & Prost, J. (2000) *Proc. Natl Acad. Sci. USA* **97**, 3183–3188.
74. Choe, Y., Magnasco, M.O. & Hudspeth, A.J. (1998) *Proc. Natl Acad. Sci. USA* **95**, 15321–15326.
75. Ruggero, M.A. (1992) *Curr. Opin. Neurobiol.* **2**, 449–456.
76. Probst, R. (1990) *Adv. Otorhinolaryngol.* **44**, 1–91.
77. Eguiluz, V.M., Ospeck, M., Choe, Y., Hudspeth, A.J. & Magnasco, M.O. (2000) *Phys. Rev. Lett.* **84**, 5232–5235.

DETECTION OF SYNCHRONY IN THE ACTIVITY OF AUDITORY NERVE FIBERS BY OCTOPUS CELLS OF THE MAMMALIAN COCHLEAR NUCLEUS

Donata Oertel*[†], Ramazan Bal*, Stephanie M. Gardner*, Philip H. Smith[‡], and Philip X. Joris[§]

Departments of *Physiology and [‡]Anatomy, University of Wisconsin Medical School, Madison, WI 53706; and [§]Division of Neurophysiology, Katholieke Universiteit Leuven, Leuven, B-3000 Belgium

The anatomical and biophysical specializations of octopus cells allow them to detect the coincident firing of groups of auditory nerve fibers and to convey the precise timing of that coincidence to their targets. Octopus cells occupy a sharply defined region of the most caudal and dorsal part of the mammalian ventral cochlear nucleus. The dendrites of octopus cells cross the bundle of auditory nerve fibers just proximal to where the fibers leave the ventral and enter the dorsal cochlear nucleus, each octopus cell spanning about one-third of the tonotopic array. Octopus cells are excited by auditory nerve fibers through the activation of rapid, calcium-permeable, α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate receptors. Synaptic responses are shaped by the unusual biophysical characteristics of octopus cells. Octopus cells have very low input resistances (about 7 M Ω), and short time constants (about 200 μ sec) as a consequence of the activation at rest of a hyper-polarization-activated mixed-cation conductance and a low-threshold, depolarization-activated potassium conductance. The low input resistance causes rapid synaptic currents to generate rapid and small synaptic potentials. Summation of small synaptic potentials from many fibers is required to bring an octopus cell to threshold. Not only does the low input resistance make individual excitatory postsynaptic potentials brief so that they must be generated within 1 msec to sum but also the voltage-sensitive conductances of octopus cells prevent firing if the activation of auditory nerve inputs is not sufficiently synchronous and depolarization is not sufficiently rapid. *In vivo* in cats, octopus cells can fire rapidly and respond with exceptionally well-timed action potentials to periodic, broadband sounds such as clicks. Thus both the anatomical specializations and the biophysical specializations make octopus cells detectors of the coincident firing of their auditory nerve fiber inputs.

Most acoustic information arrives at the brainstem of mammals through large, myelinated auditory nerve fibers that form a single, tonotopically organized pathway. In the synaptic connection of auditory nerve fibers with distinct groups of principal cells, the auditory pathway branches into multiple, parallel ascending pathways. The two groups of principal cells of the dorsal cochlear nucleus, fusiform and giant cells, project directly to the inferior colliculus. Pathways through the ventral cochlear nucleus (VCN) diverge through bushy, D stellate, T stellate, and octopus cells to take part in intermediate integrative circuits before converging again in the inferior colliculus. How these pathways contribute to the fundamental biological tasks of localizing and interpreting sounds is only partly understood. There is strong evidence that pathways through bushy cells and their targets in the medial and lateral superior olivary nuclei contribute to the localization of sound in the horizontal plane (1, 2). What integrative tasks are performed through other pathways is less well understood. The possibility has been raised that in mammals pathways through the dorsal cochlear nucleus might be involved in analyzing spectral cues for localization in the vertical plane (3). In birds, which lack a structure like the mammalian dorsal cochlear nucleus, localization in the vertical plane seems to be accomplished through homologues of T stellate cells of the mammalian VCN (4). Very little is known about how the pathways through the brainstem in vertebrates contribute to the recognition of acoustic patterns such as those in speech.

The possibility that octopus cells are involved in the recognition of natural sounds, including speech, is intriguing but untested. Octopus cells detect synchrony in the firing of groups of auditory nerve fibers, a pattern that is important for the understanding of speech. Studies from a variety of perspectives have concluded that the temporal structure in the firing of auditory nerve fibers is important in the representation of speech sounds (5, 6). Not only is phase locking important for the recognition of such fundamental features of sounds as pitch but broadband transients and gaps are critical features of consonants in speech. A second intriguing aspect of the role of octopus cells is that they are involved in largely monaural neuronal circuits. The observation that the loss of hearing in one ear does not significantly hinder speech recognition in quiet environments indicates that pattern recognition is a monaural function; octopus cells project to the contralateral ventral nucleus of the lateral lemniscus (VNLL), a nucleus that is largely monaural in most species and is likely to be involved in fundamental functions as it is present not only in mammals but also in birds and reptiles (7–10). Octopus cells also project to the superior paraolivary nucleus, a nucleus about which less is known but which is innervated mainly by the contralateral cochlear nucleus (11). The ventral nuclei of the lateral lemniscus receive input mainly from the contralateral VCN in most, but not all, species. A third intriguing observation is that there is considerable variation in the structure and in the relative proportions of inputs to the VNLL between species. It is possible that this variability reflects differences in the needs of different species in extracting biological meaning from sounds in their environment. In cats, bats, and guinea pigs the monaural, ventral lemniscal nuclei are divided into the ventral and intermediate nuclei of the lateral lemniscus on the basis of the clustering and innervation by the medial nucleus of the trapezoid body (9, 12, 13) whereas in rats and opossums two subnuclei cannot be distinguished in this area (14, 15). Interestingly the specialized region of the VNLL that is innervated by octopus

[†]To whom reprint requests should be addressed at: Department of Physiology, University of Wisconsin Medical School, 1300 University Avenue, Madison, WI 53706. E-mail: oertel@physiology.wisc.edu.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: VCN, ventral cochlear nucleus; AMPA, α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate.

cells comprises 38% of the nucleus in humans whereas it occupies only 4% of the nucleus in cats (8). Lastly, patients with auditory neuropathy whose auditory brainstem responses revealed abnormally low synchrony in auditory nerve discharge have deficits in speech recognition that are disproportionate to their hearing losses (16).

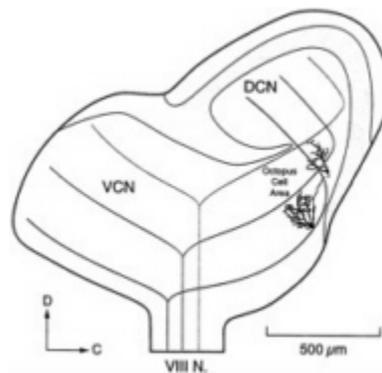


Fig. 1. Anatomical reconstructions of cell body and dendrites of two intracellularly labeled octopus cells from mice are shown on a schematic version of the cochlear nuclear complex in the parasagittal plane. The granule cell lamina (blue) separates the unlayered VCN from the layered dorsal cochlear nucleus (DCN). Octopus cells occupy an area (yellow) at the most caudal and dorsal extreme of the VCN where auditory nerve fibers are closely bundled as they cross from the VCN to the DCN. Auditory nerve fibers that encode high frequencies (light brown) terminate rostrally and those that encode low frequencies (dark brown) terminate caudally in the octopus cell area. The dendrites of octopus cells extend rostrally from the cell body. Adapted from the results of Golding *et al.* (19).

Projection of the Auditory Nerve on Octopus Cells

The tonotopic array of auditory nerve fibers is tapped systematically by octopus cell dendrites (Fig. 1). The octopus cell area occupies the most dorsal and caudal tail of the VCN where auditory nerve fibers bundle closely (17). In mice the octopus cell area has sharply defined borders and contains only octopus cells (18–20) but in other species it may be heterogeneous. Each auditory nerve fiber bifurcates at the nerve root sending one branch caudally through the posteroventral to the dorsal cochlear nucleus. Terminals of auditory nerve fibers are subtly different in the octopus cell area than in other regions of the VCN. Auditory nerve terminals that innervate bushy and stellate cells rostrally to the octopus cell area are variable in size and shape; large and small end bulbs lie intermingled with large and small boutons. In the octopus cell area, in contrast, terminals of auditory nerve fibers are uniformly small boutons. In the octopus cell area of mice, the fibers are tonotopically organized in the parasagittal plane, with fibers encoding the highest frequencies lying rostrally and those encoding the lowest frequencies caudally. The dendrites of octopus cells emanate from the rostral pole of the cell bodies so that octopus cells receive input from fibers that encode low frequencies near the cell body and from those that encode higher frequencies progressively more distally on the dendrites (17, 19–21). Octopus cell dendrites span only about one-third of the tonotopic array of auditory nerve fibers (19–21). If mice hear over a range of about 8 octaves (22), individual octopus cells would be expected to receive input from auditory nerve fibers that encode roughly between 2 and 3 octaves. In mice about 200 octopus cells (23) sample the array of about 12,000 auditory nerve fibers (24). As all auditory nerve fibers have been observed to terminate in the octopus cell area, octopus cells receive on average at least 60 inputs (25, 26). However, the number of auditory nerve inputs onto octopus cells may be several times 60 because many auditory nerve fibers probably innervate multiple octopus cells. Although most excitatory input to octopus cells is from auditory nerve fibers, in mice octopus cells also are excited through collaterals of octopus cells (19). In other species the arrangement of auditory nerve inputs on octopus cells appears to be similar but has not been investigated in as much detail. Fig. 2 shows an anatomical reconstruction of an octopus cell from a cat that was labeled by the intraxonal injection of label. The dendrites of this cell also emanate from one pole. The relationship of the tonotopic arrangement of auditory nerve fibers with respect to the dendrites of octopus cells is less clear in cats where the tonotopic axis is not aligned in the conventional planes of section.

The convergent input from a relatively large number of auditory nerve fibers is reflected in the responses of octopus cells to the activation of the auditory nerve with shocks in slices. Synaptic responses grow incrementally as more and more auditory nerve fibers are simultaneously brought to threshold with brief (0.1 msec) shocks of increasing strength (Fig. 3). Several features of synaptic responses in octopus cells are noteworthy. First, the amplitude of excitatory postsynaptic potentials varied over a wide range, from just detectable responses to weak shocks to about 30-mV responses to strong shocks; maximum amplitudes ranged between about 15 and 50 mV in different cells (19). The responses are so finely graded with shock strength that

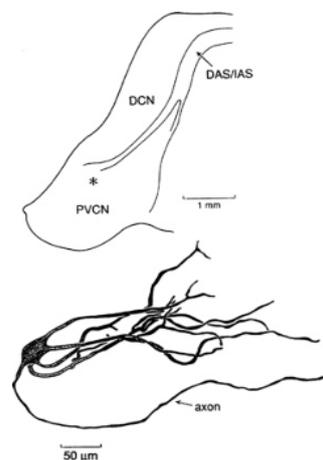


Fig. 2. Reconstruction with a camera lucida of an octopus cell in a cat that was labeled by an intraxonal injection. The location of the octopus cell body in the coronal section of the cochlear nuclear complex with respect to the posteroventral cochlear nucleus (PVCN), dorsal cochlear nucleus (DCN), dorsal acoustic stria (DAS), and intermediate acoustic stria (IAS) is indicated by * (Upper).

incremental responses from individual auditory nerve fiber inputs could not be resolved. Second, one small jump in amplitude, which was accompanied by a small action potential, was consistently detected at intermediate stimulus strengths (19) (Fig. 3, arrowhead). Responses to shocks recorded at the cell body comprised small action potentials superimposed on large synaptic potentials. Such an arrangement allows the timing of the synaptic inputs to be reflected in the timing of the action potential with precision because the relatively small action potential distorts the timing of the peak of the synaptic response only minimally. Third, over the entire range of suprathreshold responses the timing of the peaks of responses varied by only about 300 μ sec. The timing of peaks was not only consistent but also precise.

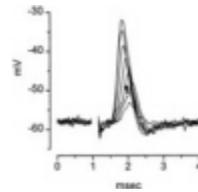


Fig. 3. Octopus cells fire in response to the coincident activation of many, but not necessarily all, of the auditory nerve fibers by which they are innervated. Seven superimposed responses are shown to shocks of the auditory nerve of 0.1 -msec duration and of varying strength (1–10 V) delivered through a pair of tungsten electrodes. Responses were recorded with a sharp microelectrode filled with 4 M potassium acetate from an octopus cell in a parasagittal slice from the cochlear nucleus of a mouse. The extracellular saline was saturated with 95% oxygen/5% carbon dioxide and contained 130 mM NaCl, 3 mM KCl, 1.3 mM MgSO₄, 2.4 mM CaCl₂, 20 mM NaHCO₃, 3 mM HEPES, 10 mM glucose, 1.2 mM KH₂PO₄, pH 7.4. Shocks produced artifacts that serve as markers of their occurrence and whose removal left a blank space in the traces. The amplitude of responses was a monotonic function of the shock strength with the weakest shocks producing the smallest responses and the strongest shocks producing the largest responses. The appearance of a small action potential, whose inflection point is marked with an arrowhead, shows where the response was just large and rapid enough to cause firing in the octopus cell. In larger responses the action potential and synaptic potential cannot be resolved. The recording was made by N.L.Golding (19).

Octopus cells detect coincidence of firing in the population of auditory nerve fibers by which they are innervated by requiring the summation of multiple synaptic inputs to reach threshold. In all recordings from octopus cells the amplitude of subthreshold synaptic responses was graded, indicating that inputs from multiple auditory nerve fibers had to sum to produce an action potential in octopus cells. The brevity of synaptic responses makes summation possible only when auditory nerve fibers are activated within about 1 msec. When auditory nerve fibers are activated with shocks in slices, activation is synchronous and summing is optimal. Under these conditions activation of roughly between one-tenth to one-third of the auditory nerve fiber inputs is required to bring octopus cells to threshold. In responses to sound *in vivo*, when auditory nerve fibers are not necessarily activated in such perfect synchrony, a larger proportion of inputs may be required to activate octopus cells.

The firing of octopus cells can follow the activation of auditory nerve fibers with temporal precision even at high rates. When shocks are delivered to octopus cells at 1/sec, the timing of the peak of the response had a standard deviation of between 20 and 40 μ sec (19). Octopus cells can respond to repeated shocks to the auditory nerve to the maximum rate at which auditory nerve fibers can be driven, about 1,000/sec. The responses of octopus cells to activation of the auditory nerve with trains of shocks to the maximum firing rate that is observed *in vivo*, 300/sec, show no depression. The timing of the peaks of responses with respect to the shock are constant. Only at unphysiological stimulation rates do the responses to octopus cells show signs of depression. Responses to the last of a 10-msec train of shocks at 714 Hz were reduced in amplitude by 25% and had a latency about 200 μ sec longer than responses to the first (19). The observed depression arises only in part from synaptic depression because the amplitude of action potentials in auditory nerve fibers are reduced at high firing rates. The ability to fire rapidly and with temporal precision also is observed in responses to sound *in vivo*. Octopus cells can respond to tones of 800 Hz at every cycle of the tone (27). *In vivo*, therefore, octopus cells have maximum firing rates that are more than double that of their auditory nerve inputs.

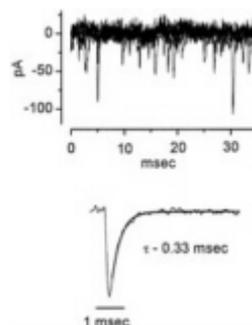


Fig. 4. Spontaneous miniature synaptic currents recorded from an octopus cell of a mouse under voltage clamp. Patch recording from an octopus cell was made in the whole-cell configuration by using a potassium gluconate-filled pipette. (Upper) Five traces are superimposed to illustrate the frequent spontaneous events when the cell was held near its resting potential at -65 mV. (Lower) Ensemble average of 113 events in the same cell. The decay of currents was well fit with a single exponential with a time constant (τ) equal to 0.33 msec. The pipette solution contained 108 mM potassium gluconate, 9 mM HEPES, 9 mM EGTA, 4.5 mM MgCl₂, 14 mM phosphocreatinine (Tris salt), 4 mM ATP (Na salt), and 0.3 mM GTP (Tris salt); pH was adjusted to 7.4 with KOH. The composition of the extracellular saline is given in the legend to Fig. 3. The results have been corrected for a junction potential of -12 mV.

The finding that the terminals of auditory nerve fibers contain high levels of glutamate suggests that glutamate is the neurotransmitter that mediates excitation (28). The glutamate released by auditory nerve fibers acts on the AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate) subtype of glutamate receptors on their targets (19). Under voltage-clamp at the resting potential numerous miniature excitatory postsynaptic currents (mEPSCs) that are sensitive to 6, 7-dinitroquinoxaline-2, 3-dione and insensitive to tetrodotoxin are observed (Fig. 4). Like other AMPA receptors in brainstem auditory neurons of mice and rats and their avian homologues, the AMPA receptors

of octopus cells are exceptionally rapid, rising from 10–90% in 0.20 ± 0.12 msec and decaying with time constants of 0.35 ± 0.16 msec at 33°C (29–31). The mEPSCs in octopus cells of mice show little sign of dendritic filtering, not only when they were recorded with patch pipettes that contained Cs^+ (31) but also when they were recorded with potassium gluconate-containing pipettes (Fig. 4). The finding that the AMPA receptors of octopus cells are blocked by the polyamine-containing wasp toxin, philanthotoxin (31), indicates that the receptors in octopus cells lack GluR2 subunits and therefore would be expected to be permeable to calcium (30, 32). It has been shown that calcium-permeable AMPA receptors have single-channel conductances that are 2–3 times larger than calcium-impermeable AMPA receptors (33–35). The low input resistance of octopus cells requires that the robust, suprathreshold synaptic potentials that are observed in octopus cells be driven by large synaptic currents. Possibly the large number of receptors required for the activation of octopus cells and their calcium permeability account for the high levels of calretinin in octopus cells (8).

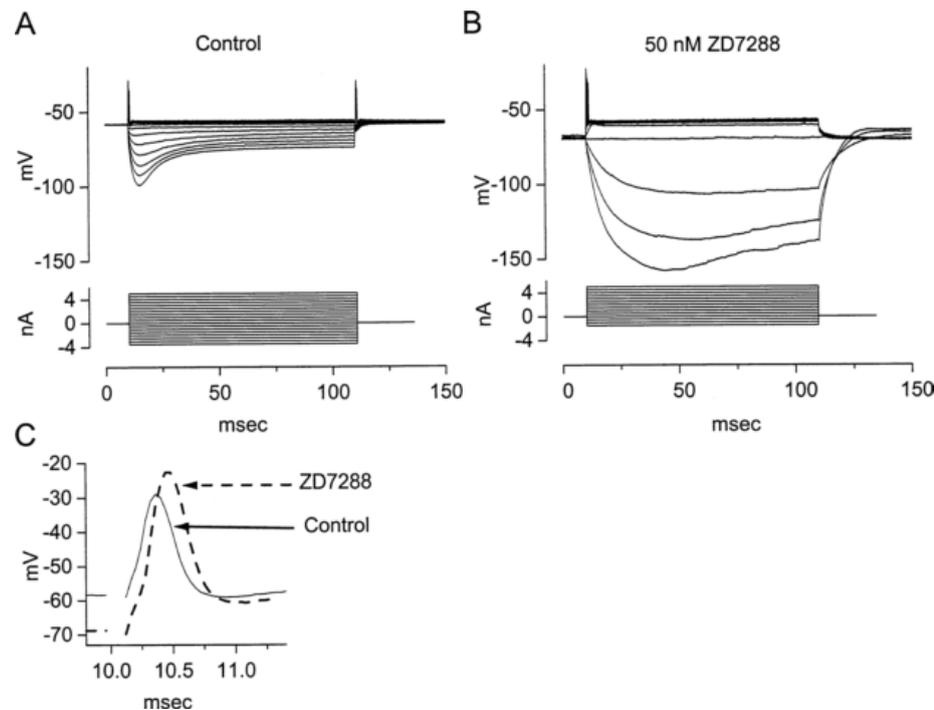


Fig. 5. (A) Polarization of an octopus cell with current pulses (-3.5 to 5 nA in 0.5 -nA steps) reveals the biophysical characteristics of the cell. Depolarizing current pulses greater than 1 nA produced small action potentials at the onset of the depolarization. After the single action potential, the octopus cell remained depolarized by a few mV. Hyperpolarizing current pulses produced transient hyperpolarizations that sagged back toward rest. (B) In the presence of 50 nM ZD7288, a blocker of g_h , the cell hyperpolarized and the input resistance in the hyperpolarizing voltage range increased. The increase in input resistance is reflected in that the current pulses produced larger and slower hyperpolarizations. The rectification in the depolarizing voltage range, reflected in the cluster of traces in responses to depolarizing currents, is not affected by ZD7288. (C) Blocking g_h shapes responses in the physiological voltage range. Expansions of the onset of responses to the largest depolarizations shown in A and B show that action potentials are taller and broader in the absence of g_h . The voltage drop across the resistance of the electrode was balanced off-line. Whole-cell patch recording from an octopus cell with solutions as in Fig. 4.

Biophysical Characteristics of Octopus Cells

The intrinsic biophysical properties of octopus cells have been studied in slices from mice. Fig. 5A illustrates the responses of octopus cells to current pulses. The resting potential of octopus cells measured with patch-clamp electrodes is 62 ± 2 mV ($n = 135$) (36). The voltage changes produced by current pulses are small both in the hyperpolarizing and depolarizing directions. When they are depolarized with current pulses greater than about 1 nA, octopus cells fire only a single, small action potential. When they are hyperpolarized, the membrane potential of octopus cells sags back toward rest after the initial hyperpolarization. The voltage-current relationships plotted from peak or steady-state levels are nonlinear (20, 36). Estimates of the input resistance, made from the slope of voltage/current relationships in the voltage range just negative to the resting potential show that octopus cells have input resistances of about 2 and 7 M Ω when measured from steady-state and peak voltage changes, respectively (20, 36).

Octopus cells have conventional regenerative currents that underlie the firing of action potentials. They generate all-or-none action potentials that are sensitive to tetrodotoxin (20). Octopus cells have exceptionally large axons (9, 11, 19, 21, 37) from which trains of action potentials have been recorded in responses to sound (38). These action potentials presumably appear small in recordings from the cell body because they are generated at an electrically distant site near the axon hillock and are attenuated as they spread back to the cell body. In the presence of α -dendrotoxin action potentials are large, suggesting that a potassium conductance provides a pathway for the leakage of depolarizing current as the action potential spreads to the cell body (M. Ferragamo and D.O., unpublished results). Octopus cells also have a weak, voltage-sensitive calcium conductance whose existence was demonstrated by blocking voltage-sensitive Na^+ and repolarizing K^+ channels and evoking broad, regenerative, calcium-sensitive action potentials (20).

Two voltage-sensitive conductances that are activated at rest dominate the biophysical properties of octopus cells. One is a hyperpolarization-activated, ZD7288-sensitive, mixed-cation conductance, g_h , and the other is a depolarization-activated, α -dendrotoxin-sensitive, low-threshold potassium conductance,

$g_{K(L)}$. Although these conductances are activated by voltage changes in the opposite direction, the voltage range of activation of the conductances overlap at the resting potential. Together these conductances set the resting potential to a level near -62 mV at which the inward current, I_h , balances the outward current, $I_{K(L)}$ (36). The experiment illustrated in Fig. 6 illustrates the balance in one cell. On average the magnitude of the inward current blocked by ZD7288 was $1,280 \pm 270$ (mean \pm SD) pA; the addition of 50 nM α -dendrotoxin in those same seven cells left an outward current of 33 ± 46 pA. The simultaneous activation of these two conductances not only makes the input resistance of octopus cells low but also endows octopus cells with biophysical characteristics that promote firing in response to synchronous inputs and prevent firing when inputs are not synchronous.

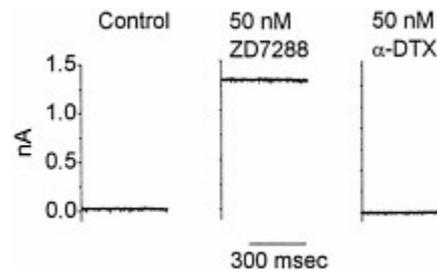


Fig. 6. At the resting potential I_h is roughly balanced by $I_{K(L)}$. The murine octopus cell was held at its resting potential, -63 mV, under voltage clamp with whole-cell patch clamp under conditions such as those described for Fig. 4. In the presence of ZD7288, a large, steady, outward current developed. Application of α -dendrotoxin blocked a current exactly equal to that which had developed in the presence of ZD7288.

The hyperpolarization-activated, mixed-cation conductance, g_h , in octopus cells resembles such conductances in other cells but is unusually large and has a half-maximal activation that is more depolarized than in most neurons (19, 20, 36). This conductance is sensitive to extracellular Cs^+ and ZD7288 (36). The reversal potential of the current through this conductance was -38 mV under normal physiological conditions and was sensitive to extracellular concentrations of both K^+ and Na^+ (20, 36). The permeability ratio P_{Na}/P_K of g_h in octopus cells was about 0.2 (36). When fully activated at hyperpolarizing potentials, the maximum g_h was 150 ± 30 nS (36). The half-maximal activation voltage, V_{half} , is unusually depolarized, lying at -65 mV. As a result of the large maximal conductance, of which a high proportion is activated at the resting potential, g_h contributes substantially to the total input conductance. At rest g_h contributed from 35 to 85 nS, with a mean of 62 nS, to the total input conductance, which was on average 149 nS (36). In the presence of ZD7288 the resting potential of octopus cells hyperpolarizes by about 10 mV (Fig. 5B). The characteristics of g_h in octopus cells indicate that this conductance is mediated by a class of ion channels that has been termed HCN (for hyperpolarization-activated and cyclic nucleotide-gated channels) (39, 40). The activation and deactivation of g_h are relatively slow with respect to the signaling of octopus cells. The fast and slow time constants of activation, τ_{fast} and τ_{slow} , were voltage-dependent with $\tau_{fast} = 44 \pm 6$ ms and $\tau_{slow} = 181 \pm 39$ ms at -77 mV and decreasing to $\tau_{fast} = 16 \pm 3$ ms and $\tau_{slow} = 84 \pm 20$ ms at -107 mV (36). Deactivation was fit with single exponentials 126 ± 15 ms at -62 mV and 178 ± 33 ms at -87 mV. Although g_h is activated by hyperpolarization, this conductance nevertheless shapes responses in the physiological, depolarizing voltage range because the rates of activation and deactivation are slow relative to the duration of synaptic potentials and action potentials. In the presence of ZD7288 action potentials rose more slowly, reached higher peaks, and were broader than under control conditions (Fig. 5C).

A depolarization-activated, low-threshold K^+ conductance also contributes to the unusual properties of octopus cells (19, 20, 36). Low-threshold K^+ conductances are prominent in many neurons in the auditory brainstem nuclei of vertebrates, causing them to fire only at the onset of current pulses (41, 42). In octopus cells, as in other brainstem neurons, this conductance is sensitive to 4-aminopyridine and α -dendrotoxin (M.Ferragamo, R.B., and D.O., unpublished results) (20). The finding that 4-aminopyridine causes the resting potential of octopus cells to depolarize indicates that the threshold of activation of the K^+ conductance is more hyperpolarized than the resting potential and identifies it as a low-threshold K^+ conductance, $g_{K(L)}$ (20). Homomeric and heteromeric channels with Kv1.1, Kv1.2, and Kv1.3 subunits have low thresholds for activation (43, 44). Immunocytochemical labeling for α subunits of K^+ channels of the Kv1 family suggests that potassium channels of this family may underlie $g_{K(L)}$. Potassium channel α subunits Kv1.1 and Kv1.2 (45) have been shown to be strongly expressed in the octopus cell area.

Less is known about other K^+ conductances in octopus cells. Under conditions when $g_{K(L)}$ was blocked with 4-aminopyridine or α -dendrotoxin, action potentials repolarized slowly (20). Immunolabeling for high-threshold, Kv3.1 potassium channels has been detected in the cell bodies of octopus cells (46).

Despite the large conductances that are active at rest, three experimental observations suggest that dendritic filtering is surprisingly low in octopus cells. The first is that recordings of miniature synaptic currents in octopus cells showed no sign of dendritic filtering (31). Octopus cells receive input from auditory nerve fibers on dendrites (19). Dendritic filtering would be expected to produce a positive correlation between rise and fall times and a negative correlation between rise time and amplitude but none was observed (31). Miniature synaptic currents were uniformly rapid not only when intracellular Cs^+ in the recording pipettes were used to block leakage in octopus cells but also when pipettes contained potassium gluconate. The second observation indicative of isopotentiality and lack of filtering in octopus cell dendrites is that I_h recorded under voltage clamp was well-behaved; chord conductances converged at a single point under a wide range of conditions. Third, in the study of g_h the reversal potential of I_h , a mixed-cation current, was measured when the extracellular Na^+ and K^+ concentrations, and therefore the reversal potential, was varied. To test whether the reversal potentials measured under these conditions were consistent, the relative permeabilities to Na^+ and K^+ were calculated and compared and found not to be statistically different from one another (36). Whether the low dendritic filtering results primarily from the large size of dendrites or from a favorable spatial distribution of ion channels is not clear.

The interplay of conductances gives octopus cells unusual biophysical properties. When depolarized with a steady pulse of current, octopus cells fire only once at the onset (Fig. 5); in no octopus cell have multiple action potentials ever been observed in responses to depolarizing current pulses. The presence of two voltage-sensitive conductances at rest also makes the firing of octopus cells sensitive to the rate at which they are depolarized. Octopus cells fire when they are depolarized rapidly but fail to fire when they are depolarized slowly (M.Ferragamo and D.O., unpublished results) (47). The finding that octopus cells fire only once in response to long depolarizations does not preclude their being able to fire rapidly. A train of current pulses presented at 1,000/sec drives action potentials in an octopus cell with every pulse (Fig. 7). Not surprisingly, the first action potential is larger than the later ones, which rise from the undershoot of the

preceding action potentials. These observations raise the question to what extent octopus cells are refractory after the first action potential. The experiment illustrated in Fig. 8 shows that octopus cells can be induced to fire even when they are steadily depolarized. This octopus cell was depolarized with a current pulse of 2 nA and then the current was increased in two steps. With each increase the octopus cell fired an action potential. The later action potentials were smaller, when measured from the inflection to the peak, than the first, presumably because the regenerative inward current had to counter the larger, steady outward current. The large after-hyperpolarization that followed the offset of the current reflects the deactivation of the potassium conductance that had been activated by the previous depolarization.

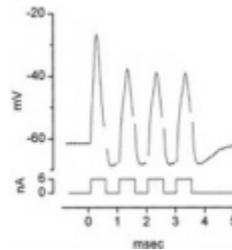


Fig. 7. Octopus cells can fire rapidly. A train of depolarizing current pulses presented at 1,000/sec evoked action potentials at every pulse. The voltage drop across the resistance of the electrode was balanced off-line; removal of transient artifacts left brief gaps in the trace. Whole-cell patch recording from an octopus cell of a mouse with solutions as in Fig. 4.

Responses to Sound

Few reports have been published of the responses to sound of neurons that are identified with some certainty of being octopus cells. Godfrey *et al.* (48) concluded that octopus cells in cats respond to tones >2 kHz with sharply timed action potentials at the onset. This conclusion was confirmed by later studies allowing octopus cells to be identified by their onset responses (27, 37, 38, 49, 50). Recordings *in vivo* indicate that the anatomical and biophysical features of octopus cells that have been revealed *in vitro* are correlated with the ability of neurons to encode temporal features of acoustic stimuli with greater precision than their auditory nerve inputs and with greater precision than other groups of neurons in the cochlear nuclei. Consistent with the observation that octopus cells are innervated by many auditory nerve fibers and require the synchronous activation of a substantial fraction of those inputs, octopus cells are broadly tuned and have high thresholds to pure tones (27, 48, 51). At high intensities a broad range of auditory nerve fibers can respond to tones of frequencies less than about 2 kHz, with discharges that are locked to a particular phase of each stimulus cycle. Octopus cells can respond to such tones with a single well-timed spike at every stimulus cycle for frequencies up to 800 Hz, firing at rates that are unprecedented in the central nervous system. They respond to tones above about 2 kHz with a single action potential at the onset of the tone, presumably because it is only at stimulus onset that the firing of auditory nerve inputs fire in sufficient synchrony to drive octopus cells. Octopus cells also respond to broadband transients such as clicks with exceptionally well-timed action potentials. Of all cells in the cochlear nucleus, the octopus cells show the strongest synchronization to amplitude-modulated stimuli (49) and to the fundamental frequency of simple speech-like sounds (50). Not only is the precision in the timing of firing

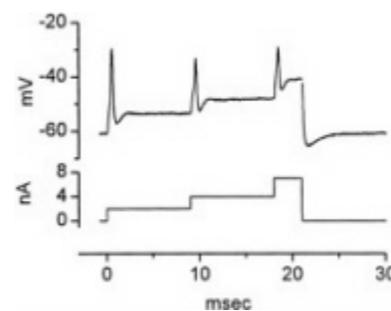


Fig. 8. When steadily depolarized, octopus cells are in a relative, but not absolute, refractory period. This octopus cell was depolarized with increasing current steps. Depolarization with a current pulse of 2 nA produced an action potential. Further step depolarizations from 2 to 4 nA and from 4 to 7 nA caused the octopus cell to fire again. Action potentials that were evoked by superimposed step depolarizations were smaller than the initial action potential. At the offset of the current pulse, the octopus cell undershot the resting potential. The voltage drop across the resistance of the electrode was balanced off-line and transient artifacts were made blank. Whole-cell patch recording from a murine octopus cell with solutions as in Fig. 4.

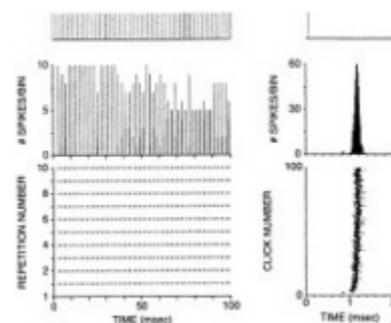


Fig. 9. Response of an octopus cell to clicks in a cat *in vivo*. (Left) Responses to 10 repetitions of a train of acoustic clicks (20- μ sec duration) spaced at 2-msec (500 Hz) intervals. (Top) Trace shows the timing of the click stimuli that were presented 30 dB above threshold, (Middle) the poststimulus time histogram (0.2-msec binwidth), and (Bottom) dot rasters to 10 click trains with each symbol representing one action potential. (Right) Responses on an expanded time scale as a function of the 2-msec period of the stimulus. (Top) Trace shows position of the click in the period, (Middle) the period histogram using 8- μ sec binwidths, and (Bottom) dot rasters ordered by click number (response to first click on bottom, response to 100th click on top). The threshold of responses to tones at the characteristic frequency, 9 kHz was 52 dB sound pressure level.

remarkable but these cells also show sharp tuning to modulation frequency in terms of average firing rate.

Responses to clicks in one octopus cell from a cat are illustrated in Fig. 9. The recording was made intracellularly from the axon with a dye-filled pipette, making it possible to reconstruct the cell from which the recording was made after the end of the experiment (Fig. 2). The cell responded to a train of clicks at 500 Hz with trains of action potentials whose timing followed the stimulus with precision (Fig. 9). The dot raster in the lower left panel, showing the spike response to 10 repetitions of a 100-ms train of clicks, is regular. The histogram in the left, middle panel shows that the spikes fall into either one or two 0.2-ms bins. The timing of the firing is illustrated in greater resolution on the right as a function of the 2-ms period of the stimulus. The timing of firing of individual action potentials is shown in the dot raster plot at the bottom and is shown as a histogram with 8- μ sec bins in the middle. These records show that the jitter in the timing of firing of octopus cells is less than 200 μ sec.

Despite the large size of octopus cells, *in vivo* recordings have proven to be surprisingly difficult to make. The biophysical properties of octopus cells perhaps can account for that difficulty. The finding that action potentials recorded at the cell body are small suggests that extracellular currents associated with those action potentials are also small and difficult to record. Although the action potentials associated with the axons are easier to record, the axons themselves are not easy to reach (38).

The ideas and conclusions summarized here reflect the thoughts and efforts of many people whose substantial contributions are a pleasure to acknowledge. When Shu Hui Wu first revealed the morphology of octopus cells, we had not yet appreciated how fascinating they were. It was the findings of Robert Wickesberg and Donna Whitlon that drew our attention to the octopus cell area. Nace Golding, Don Robertson, and Michael Ferragamo then made critical observations that serve as the basis of our recent conclusions. This work depended on the support by National Institutes of Health Grants DC00176 and DC00116.

1. Joris, P.X., Smith, P.H. & Yin, T.C. (1998) *Neuron* **21**, 1235–1238.
2. Oertel, D. (1999) *Annu. Rev. Physiol.* **61**, 497–519.
3. Spirou, G.A., Davis, K.A., Nelken, I. & Young, E.D. (1999) *J. Neurophysiol.* **82**, 648–663.
4. Takahashi, T., Moiseff, A. & Konishi, M. (1984) *J. Neurosci.* **4**, 1781–1786.
5. Young, E.D. & Sachs, M.B. (1979) *J. Acoust. Soc. Am.* **66**, 1381–1403.
6. Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J. & Ekelid, M. (1995) *Science* **270**, 303–304.
7. Vater, M. & Feng, A.S. (1990) *J. Comp. Neurol.* **292**, 373–395.
8. Adams, J.C. (1997) *Aud. Neurosci.* **3**, 335–350.
9. Schofield, B.R. & Cant, N.B. (1997) *J. Comp. Neurol.* **379**, 363–385.
10. Carr, C.E. & Code, R.A. (2000) in *The Central Auditory System of Reptiles and Birds*, eds. Dooling, R., Popper, A.N. & Fay, R.R., Springer Handbook of Auditory Research 13 (Springer, New York), pp. 197–248.
11. Schofield, B.R. (1995) *J. Comp. Neurol.* **360**, 135–149.
12. Glendenning, K.K., Brunso-Bechtold, J.K., Thompson, G.C. & Masterton, R.B. (1981) *J. Comp. Neurol.* **197**, 673–703.
13. Covey, E. & Casseday, J.H. (1986) *J. Neurosci.* **6**, 2926–2940.
14. Willard, F.H. & Martin, G.F. (1983) *Neuroscience* **10**, 1203–1232.
15. Merchan, M.A. & Berbel, P. (1996) *J. Comp. Neurol.* **372**, 245–263.
16. Zeng, F.G., Oba, S., Garde, S., Sininger, Y. & Starr, A. (1999) *NeuroReport* **10**, 3429–3435.
17. Osen, K.K. (1969) *Acta Otolaryngol (Stockh.)* **67**, 352–359.
18. Wickesberg, R.E., Whitlon, D.S. & Oertel, D. (1994) *J. Comp. Neurol.* **339**, 311–327.
19. Golding, N.L., Robertson, D. & Oertel, D. (1995) *J. Neurosci.* **15**, 3138–3153.
20. Golding, N.L., Ferragamo, M. & Oertel, D. (1999) *J. Neurosci.* **19**, 2897–2905.
21. Oertel, D., Wu, S.H., Garb, M.W. & Dizack, C. (1990) *J. Comp. Neurol.* **295**, 136–154.
22. Ehret, G. (1983) in *Psychoacoustics*, ed. Willott, J.F. (Thomas, Springfield), pp. 13–56.
23. Willott, J.F. & Bross, L.S. (1990) *J. Comp. Neurol.* **300**, 61–81.
24. Ehret, G. (1979) *J. Comp. Neurol.* **183**, 73–88.
25. Lorente de No, R. (1933) *Laryngoscope* **43**, 327–350.
26. Brown, M.C. & Ledwith, J.V. (1990) *Hear. Res.* **49**, 105–118.
27. Rhode, W.S. & Smith, P.H. (1986) *J. Neurophysiol.* **56**, 261–286.
28. Hackney, C.M., Osen, K.K., Ottersen, O.P., Storm-Mathisen, J. & Manjaly, G. (1996) *Eur. J. Neurosci.* **8**, 79–91.
29. Raman, I.M. & Trussell, L.O. (1992) *Neuron* **9**, 173–186.
30. Geiger, J.R., Melcher, T., Koh, D.S., Sakmann, B., Seeburg, P.H., Jonas, P. & Monyer, H. (1995) *Neuron* **15**, 193–204.
31. Gardner, S.M., Trussell, L.O. & Oertel, D. (1999) *J. Neurosci.* **19**, 8721–8729.
32. Jonas, P. & Burnashev, N. (1995) *Neuron* **15**, 987–990.
33. Koh, D.S., Geiger, J.R., Jonas, P. & Sakmann, B. (1995) *J. Physiol (London)* **485**, 383–402.
34. Raman, I.M. & Trussell, L.O. (1995) *Biophys. J.* **68**, 137–146.
35. Swanson, G.T., Kamboj, S.K. & Cull-Candy, S.G. (1997) *J. Neurosci.* **17**, 58–69.
36. Bal, R. & Oertel, D. (2000) *J. Neurophysiol.* **84**, 806–817.
37. Rhode, W.S., Oertel, D. & Smith, P.H. (1983) *J. Comp. Neurol.* **213**, 448–463.
38. Smith, P.H., Joris, P.X., Banks, M.I. & Yin, T.C.T. (1993) in *The Mammalian Cochlear Nuclei: Organization and Function*, eds. Merchan, M.A., Juiz, J.M., Godfrey, D.A. & Mugnaini, E. (Plenum, New York), pp. 349–360.
39. Clapham, D.E. (1998) *Neuron* **21**, 5–7.
40. Seifert, R., Scholten, A., Gauss, R., Mincheva, A., Lichter, P. & Kaupp, U.B. (1999) *Proc. Natl Acad. Sci. USA* **96**, 9391–9396.
41. Oertel, D. (1997) *Neuron* **19**, 959–962.
42. Trussell, L.O. (1999) *Annu. Rev. Physiol.* **61**, 477–496.
43. Stuhmer, W., Ruppertsberg, J.P., Schroter, K.H., Sakmann, B., Stocker, M., Giese, K.P., Perschke, A., Baumann, A. & Pongs, O. (1989) *EMBO J.* **8**, 3235–3244.
44. Hopkins, W.F., Allen, M.L., Houamed, K.M. & Tempel, B.L. (1994) *Pflügers Arch. Eur. J. Physiol.* **428**, 382–390.
45. Wang, H., Kunkel, D.D., Schwartzkroin, P.A. & Tempel, B.L. (1994) *J. Neurosci.* **14**, 4588–4599.
46. Perney, T.M. & Kaczmarek, L.K. (1997) *J. Comp. Neurol.* **386**, 178–202.
47. Cai, Y., McGee, J. & Walsh, E.J. (2000) *J. Neurophysiol.* **83**, 301–314.
48. Godfrey, D.A., Kiang, N.Y.S. & Norris, B.E. (1975) *J. Comp. Neurol.* **162**, 247–268.
49. Rhode, W.S. (1994) *Hear. Res.* **77**, 43–68.
50. Rhode, W.S. (1998) *Hear. Res.* **117**, 39–56.
51. Cai, Y., Walsh, E.J. & McGee, J. (1997) *J. Neurophysiol.* **78**, 872–883.

LINEAR AND NONLINEAR PATHWAYS OF SPECTRAL INFORMATION TRANSMISSION IN THE COCHLEAR NUCLEUS

Jane J. Yu and Eric D. Young*

Department of Biomedical Engineering and Center for Hearing Sciences, Johns Hopkins University, 720 Rutland Avenue, Baltimore, MD 21205

At the level of the cochlear nucleus (CN), the auditory pathway divides into several parallel circuits, each of which provides a different representation of the acoustic signal. Here, the representation of the power spectrum of an acoustic signal is analyzed for two CN principal cells—chopper neurons of the ventral CN and type IV neurons of the dorsal CN. The analysis is based on a weighting function model that relates the discharge rate of a neuron to first- and second-order transformations of the power spectrum. In chopper neurons, the transformation of spectral level into rate is a linear (i.e., first-order) or nearly linear function. This transformation is a predominantly excitatory process involving multiple frequency components, centered in a narrow frequency range about best frequency, that usually are processed independently of each other. In contrast, type IV neurons encode spectral information linearly only near threshold. At higher stimulus levels, these neurons are strongly inhibited by spectral notches, a behavior that cannot be explained by level transformations of first- or second-order. Type IV weighting functions reveal complex excitatory and inhibitory interactions that involve frequency components spanning a wider range than that seen in choppers. These findings suggest that chopper and type IV neurons form parallel pathways of spectral information transmission that are governed by two different mechanisms. Although choppers use a predominantly linear mechanism to transmit tonotopic representations of spectra, type IV neurons use highly nonlinear processes to signal the presence of wide-band spectral features.

As sensory systems are studied more completely, it is apparent that they are composed of multiple parallel subsystems (for example, see refs. 1–4). The physiological characteristics of these segregated pathways are often quite different, leading to the assumption that each subsystem plays a specialized role in the processing of sensory information. The cochlear nucleus (CN), the termination zone of auditory nerve fibers, is an ideal structure in which to investigate the parallel processing of information within a sensory system. Not only does the CN contain at least seven different principal cell types—each with distinct morphological and physiological properties—but these distinct neuron classes project to auditory nuclei in anatomically segregated pathways (5, 6). These findings suggest that the CN is the origin of multiple functionally distinct subsystems of auditory information processing.

The parallel organization of sensory systems suggests that different neural populations may be responsible for processing different aspects of a sensory stimulus. Distinct neuron classes of the CN likely exist to decompose complex natural sounds into more simple information-bearing elements—forms that are necessary for subsequent processing in other auditory nuclei. What acoustic information is contained within each of these elements and by what mechanism do the auditory neurons compute this information? In this paper we address these questions by considering only information that is present in the power spectra of sounds. Sound spectra, distribution plots of the energy content in a stimulus across frequency, have been shown to convey information necessary for the identification and localization of behaviorally important acoustic stimuli. For example, the identity of a speech vowel appears to be determined by its formant frequencies, the frequencies at which there are peaks of energy in the power spectrum (7). We limit our discussion in this paper to the behavior of chopper and type IV neurons in the ventral CN (VCN) and dorsal CN (DCN), respectively. Both neuron types have been implicated in aspects of the representation of the frequency spectra of stimuli (8, 9). We will show that these two principal neurons of the CN compute spectral information about the acoustic environment by using different mechanisms.

Importance of Spectral Shape in Sound Localization

Spectral shapes of acoustic stimuli provide cues that are necessary for accurate sound localization in cats and humans (10, 11). Spectral sound localization cues are produced by the direction-dependence of sound propagation through the external ear. This dependence is captured by the head-related transfer function (HRTF), a measure of the ratio of the sound pressure near the eardrum to the sound pressure in free field (12, 13). Two examples are shown in Fig. 1 (14). Flat spectrum broadband noise presented in free field at two different spatial locations will be modified by two different HRTFs, like those shown in Fig. 1. As a result, spectra at the eardrum are different and can be used to identify the sound source directions. In the midfrequency region between 8 and 18 kHz, HRTFs commonly exhibit broad spectral notches (at 13.6 kHz for EO and just below 10 kHz for E2) with center frequency positions that vary systematically with source position in the frontal field. The variation of first notch position with azimuth and elevation suggests that notch positions occurring at each ear are in principle sufficient to provide a unique determinant of sound location (14). The complex spectral patterns seen at high frequencies (above 18 kHz) also change rapidly as the sound source is moved. These high-frequency cues have been shown to provide cats with the ability to discriminate between two source locations (15); however, in cats, the midfrequency notch appears to be necessary for localizing a single sound source (11).

Representation of Spectral Shape Across Neural Populations

At threshold, auditory neurons respond most strongly to a single frequency called the best frequency (BF). At higher sound levels,

*To whom reprint requests should be addressed at: 505 Traylor Building, Johns Hopkins University, 720 Rutland Avenue, Baltimore, MD 21205. E-mail: eyoung@bme.jhu.edu.

This paper was presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: BF, best frequency; CN, cochlear nucleus; DCN, dorsal CN; VCN, ventral CN; RSS, random spectral shape; HRTF, head-related transfer function.

they respond primarily to frequency regions surrounding BF. Auditory nuclei are tonotopically organized, meaning that BF values of neurons at sequential anatomical positions differ in a systematic, sequential manner (16). For example, more ventral areas of the DCN house neurons with progressively lower BF values than those in more dorsal areas (17). Tonotopic organization can be viewed as a physiological correlate in the brain to the frequency axis of the sound power spectrum.

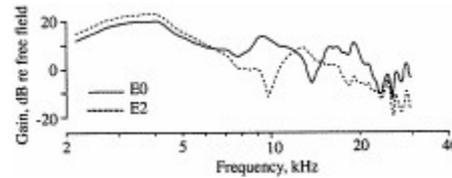


Fig. 1. Examples of HRTFs for the external ear of the cat. Plot shows gain versus frequency for sound propagation from free field to a point near a cat's eardrum (14) for two directions in space. Gains were computed as the ratio of the sound pressure measured near the eardrum divided by the sound pressure in free field, i.e., with the cat removed. E0 was 15° to the right of midline (the recording microphone was in the right ear) and 30° above the horizon. E2 was also 15° to the right, but was 15° below the horizon. Test stimuli in later figures were generated by filtering broadband noise with these and similar gain functions, thereby producing stimuli whose spectra had the shape of these gain functions (for details, see ref. 18). Earphone delivery of these stimuli simulates presentation of a broadband noise from the corresponding directions in free field.

The BF tuning of auditory neurons is the basis for two different mechanisms by which sound spectra can be represented across a neural population. In response to an auditory stimulus, an array of neurons may produce a tonotopically organized distribution of average rates that are proportional to spectral shape (e.g., ref. 18). This tonotopic representation must involve narrowly tuned neurons that produce discharge rates that are monotonically related to the spectral energy near BF. Alternatively, auditory neurons may provide a tonotopic representation of stimulus features. In this case, the responses of neurons strongly depend on the presence of wide-band spectral shapes within their receptive fields. Previous studies have shown that VCN chopper neurons possess the physiological characteristics to provide a tonotopic representation of a sound spectrum (8). In contrast, DCN type IV neurons are likely to be feature detectors, exhibiting strong rate sensitivity to the presence of wide-band spectral notches at BF (9, 19). The current study corroborates these earlier findings by applying techniques of system identification to explore the two mechanisms for the rate-encoding of spectral information.

A System Identification Approach to Understanding Spectral Processing in Auditory Neurons

Neurons producing the spectral shape representations described above are assumed here to pass spectral information to other areas of the brain in the form of average discharge rate. The transformation of spectral information into rate is assumed to be given by the following stimulus-response function:

$$r = R_0 + \sum_{j=1}^{j_2} w_j S(f_j) + \sum_{j=1}^{j_2} \sum_{k=j}^{j_2} w_{jk} S(f_j) S(f_k), \quad [1]$$

where discharge rate r is a weighted sum of energy at different frequency components of the stimulus. The frequency axis is divided into bins of 1/8-octave width with values $S(f_j)$ equal to the energy in one such bin. The spectrum S is expressed as dB relative to the average bin power across frequency. In the first sum of Eq. 1, the first-order weight, w_j , quantifies the contribution of the bin with center frequency, f_j , to the overall rate response. In the second sum, the second-order weight, w_{jk} , captures either quadratic dependence of rate on stimulus energy at a single frequency, f_j , or interactions between energy at frequency pairs f_j and f_k . R_0 is the rate response to a sound with constant spectral level of 0 dB re: the average value. The form of this model is motivated by the fact that, in response to small fluctuations in level of a fixed broadband stimulus, auditory nerve fibers and VCN neurons exhibit discharge rates that are approximately linearly related to spectral level at BF (20–22). Eq. 1 is a small-signal linearization of a nonlinear system that is valid over a restricted range of spectral levels $S(f)$. Although the range of the approximation has not been investigated, it seems to be large enough to encompass the range of spectral levels seen in HRTFs, where levels $S(f_j)$ vary by less than 10–20 dB from an average value. Similar models have been used previously (e.g., refs. 23–25).

The system parameters R_0 , w_j , and w_{jk} can be deduced by presenting stimuli with predetermined spectral shapes and recording observed discharge rates. Responses to spectra with a constant level of 0 dB are averaged to obtain an estimate of R_0 . First- and second-order weights then are estimated by recording responses to stimuli with random spectral shapes (RSS) and solving Eq. 1 by using least squares (normal equations, ref. 26). Spectra of three of the RSS stimuli are shown in the bottom three plots of Fig. 2B. The amplitudes of the spectra in each frequency bin are pseudorandom with mean 0 dB and standard deviation 12 dB. The waveforms of the stimulus set are generated by first discretizing spectra into tones that are logarithmically spaced at frequency steps of 1/64 octave. Each of the frequency bins are 1/8-octave wide and contain eight of these tones, each with identical amplitude. Sinusoidal tones with random phase are summed for each spectrum to generate the RSS stimuli. Because the model is valid only over a limited level range, model parameters are computed for different overall stimulus levels at 10-dB intervals.

The first-order weights shown in Fig. 2C were obtained from the rate responses to 192 RSS stimuli. First-order weights for two different neurons, a VCN chopper (Fig. 2C Left) and DCN type IV (Fig. 2C Right), are plotted as a function of frequency in terms of octaves relative to BF. The mean and standard deviation of the weights are determined by using a bootstrap procedure (27) in which weights are computed repeatedly by sampling with replacement from the set of 192 stimulus-rate pairs. As many as 192 weights (total number of first- and second-order weights) can be computed by using the least-squares fitting method; however, only a subset of these are useful in characterizing a neuron. Three criteria are used to estimate a useful range of first-order weights ($[j_1, j_2]$ in Eq. 1). First, a weight is considered significant if it is at least one standard deviation (determined from the bootstrap calculations) away from zero. Second, weighting functions are assumed to be continuous. That is, weights near zero that are located between two significant weights are included in the range. Third, weights considered significant must improve the accuracy with which Eq. 1 can predict responses to novel stimuli (described below). The filled circles in Fig. 2C indicate weights that are considered significant under these criteria. The range of weights $[j_1, j_2]$ defines the bandwidth of the neuron. Second-order weights are computed over the same or wider bandwidth to ensure that the higher-order model includes all contributing frequency bins.

If Eq. 1 truly describes the transformation of spectral level into rate, then estimated parameters should accurately predict responses to novel stimuli. A set of broadband noise stimuli filtered by HRTFs is used to test the model and study the representation of behaviorally relevant spectral cues. The 100

HRTFs chosen for the stimulus set were recorded from locations in the frontal field of a cat at 15° intervals in azimuth and 7.5° intervals in elevation (14). Parameters estimated by using RSS stimuli and found to be significant are used to predict responses to HRTF stimuli of matched average level. The quality of the model fit is quantified by a measure Q given by

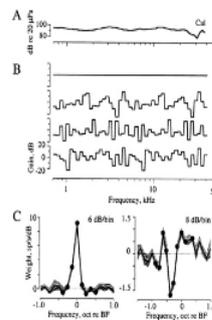


Fig. 2. (A) The acoustic calibration from one experiment. This represents the sound pressure level near the eardrum across a range of frequencies, given fixed electrical signal amplitude (0-dB attenuation) at the input of the earphone. Broadband noise with a flat spectrum at the input to the earphone will have this spectrum. All stimuli presented during an experiment are modified by the shape of this calibration, (B) The spectral envelopes of four of a set of 200 stimuli (each 5.875 octaves wide) that were used to compute the unknown system parameters of Eq. 1. The first (flat) spectrum is used to estimate R_o . The remaining three spectra are examples of RSS stimuli. Although spectra are periodic along the frequency axis, a single period is wide enough to encompass the receptive field of the neurons studied. Ordinate values of these spectra are given in dB re: the average sound level in the stimulus. As explained in the text, overall sound levels were varied systematically during the experiment. (C) Examples of first-order weighting functions for two neurons whose response maps are shown in Fig. 3. Mean weights (bold lines) and ± 1 standard deviation (gray shaded region) are computed by using a bootstrap procedure. The weights indicated by $\hat{\cdot}$ are significantly different from zero (see text). Weights were computed from RSS stimuli presented at the average spectral levels given in the legends.

$$Q = \frac{1}{1 + \frac{\sum_{i=1}^n (r_i - \hat{r}_i)^2}{\sum_{i=1}^n (\hat{r}_i - \bar{\hat{r}})^2}}, \quad [2]$$

where r_i are the experimental rates, \hat{r}_i are the rates predicted by the model, and $\bar{\hat{r}}$ is the mean of the \hat{r}_i . Q varies from 0 to 1 where 0 indicates a poor fit and 1 is obtained for a perfect fit.

Responses of Chopper and Type IV Neurons to Spectral Notches

The results shown in this paper were obtained in the CN of unanesthetized, decerebrate cats by using standard extracellular recording techniques. Stimuli were applied through a calibrated closed acoustic system. An example of an acoustic calibration is shown in the top plot of Fig. 2A. Details of the animal preparation can be found in recent publications (28–30). All procedures were approved by the Johns Hopkins Animal Care and Use Committee.

Fig. 3 illustrates how chopper and type IV neurons differ in their physiological response properties. Fig. 3A1 and B1 are response maps that show excitatory and inhibitory frequency regions in the receptive fields of a chopper and a type IV neuron at multiple sound levels. Each subplot within a response map is constructed from average rate responses to a set of 100 tone bursts whose frequencies are logarithmically interpolated across

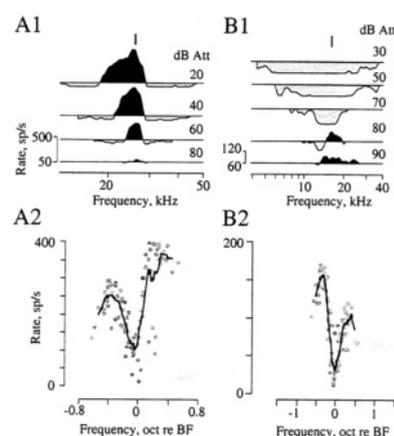


Fig. 3. (A1) Response map of a VCN chopper neuron. The V-shaped excitatory area is centered on BF (25.5 kHz; dark gray vertical line) and flanked by inhibitory areas at higher and lower frequencies. Each subplot shows average rate responses to a 200-ms tone sweep over a range of frequencies at fixed attenuation. Attenuation values are given to the right of the plots. Zero-decibel attenuation corresponds roughly to 90-to 100-dB re: 20 μ Pa; however, the actual sound pressure level varies slightly across frequency because of the acoustic calibration (see the example in Fig. 2A). The rate scale shown at the lower left applies to the 80-dB attenuation subplot. Subsequent plots are shifted vertically but use the same scale factor. Horizontal lines indicate the spontaneous rate (same at all levels). (A2) Discharge rate of the same chopper in response to HRTF-filtered noise stimuli (400-ms duration, ≈ 40 dB re: threshold) containing salient midfrequency spectral notches. Each observed rate (gray filled circle) is plotted against the frequency at which the notch occurs relative to BF. The solid black curve is a 1/8-octave smooth of the points. The horizontal gray bar indicates the spontaneous rate. This chopper encodes notch position by producing a rate minimum when the notch is centered on BF. (B1) Response map of a DCN type IV neuron (BF=15.9 kHz) characterized by a mixture of excitatory and inhibitory areas (see text). (B2) Notch responses for the same type IV neuron at ≈ 30 dB re: threshold. Although the response is qualitatively similar in shape to that in A2, this neuron is inhibited when the notch is centered on BF. Note that the logarithmic frequency axes in parts 1 and 2 are the same for both A and B.

a range spanning the response area of the neuron. Regions of rate increase above spontaneous activity are excitatory areas, whereas regions of decrease are inhibitory areas. The VCN chopper response map in Fig. 3A1 is typical of auditory neurons in that it has a V-shaped excitatory area centered on BF (16). This particular chopper neuron also exhibits inhibitory side bands. In contrast, the DCN type IV response map in Fig. 3B1 exhibits more complex patterns of inhibition. At low levels there is a small excitatory area that represents the tip of the tuning curve of the cell's excitatory inputs. At higher levels, a substantial inhibitory area appears that is centered on or just below BF. This central inhibitory area is likely to be the result of glycinergic input from DCN interneurons (vertical cells; ref. 31). At frequencies away from BF, type IV response maps vary from neuron to neuron. In general, type IV neurons exhibit a wideband inhibitory area that extends both above and below BF (30). This inhibitory area is likely to be derived from two inputs: the first being a glycinergic D-multipolar or radiate neuron in VCN, and the second being a GABAergic input whose source has not been identified. Detailed features of type IV response maps have been described elsewhere (19, 28).

Chopper and type IV neurons also differ in the way that they respond to spectral notches. Fig. 3A2 shows responses of the same chopper neuron to HRTF stimuli with spectral notches located at different frequency positions relative to the neuron's BF. Each point indicates the average discharge rate of the chopper in response to an HRTF stimulus, plotted as a function of the stimulus notch frequency relative to BF. The result shows a single-neuron analog of a tonotopic population representation in that the rate response is at a minimum when the notch is centered on BF and increases as the notch moves away from BF. In this chopper, as in auditory nerve fibers (32), the rate minimum occurs because of the reduction in sound power within the excitatory area when the notch is centered on BF. The response of the DCN type IV neuron to a spectral notch (Fig. 3B2) is qualitatively similar to that of the chopper neuron in that there is a minimum discharge rate when the notch is centered on BF. However, there are two important differences. First, the notch response goes below spontaneous rate in the type IV neuron when the notch is centered on BF. This inhibitory response does not occur in the chopper neuron. Second, notch responses are qualitatively consistent with the chopper response map in that centering the notch on BF removes energy from an excitatory response area. This should produce a reduction in discharge rate, although not necessarily an inhibitory response. In contrast, the type IV neuron gives predominantly inhibitory responses to tone energy near BF at most sound levels. Centering a notch on BF therefore should produce an excitatory response instead of the inhibitory one actually observed. The explanation for this contradiction is the dual inhibitory nature of the DCN circuit (28, 30). Type IV neurons are inhibited by vertical cells in response to tones and by D-multipolar and perhaps also GABAergic neurons in response to noise-notch stimuli.

For Chopper and Near-Threshold Type IV Neurons, the Model Accurately Predicts Responses to HRTF Stimuli

Near threshold, chopper and type IV neurons respond to noise spectra in a similar manner. Weighting functions determined within 20 dB of threshold for a chopper (Fig. 4A1) and a type IV (Fig. 4B1) neuron both are triangular with peak excitatory (positive) weights that occur within 1/8 octave of BF. Negative values in the chopper function indicate that this neuron is also weakly inhibited at frequencies below BF. First-order weights that were found to be significant were used to predict responses to notch stimuli. The quality of first-order predictions for the chopper ($Q_1=0.46$; Fig. 4A3) and type IV ($Q_1=0.42$; Fig. 4B3) neurons are moderate, suggesting that both behave reasonably linearly in the regime near threshold (see also ref. 29). However, in both cases, when the notch is located above BF, predicted rates undershoot actual rates.

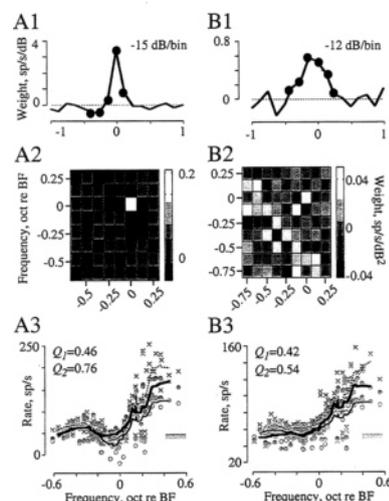


Fig. 4. Notch response predictions at sound levels near threshold. (A1) First-order weighting function for a chopper neuron (BF=25.5 kHz; same neuron in Fig. 3A), plotted as in Fig. 2. Data were taken ≈ 20 dB above threshold. (A2) Second-order weighting functions for the same chopper. Color scale bar is shown at right. Note that the plot is symmetric about the diagonal, i.e., $W_{jk} S(f_j) S(f_k)$ is the same as $W_{kj} S(f_k) S(f_j)$ and each such pair of terms is included in Eq. 1 only once. The largest weight corresponds to a quadratic term involving the frequency bin at BF (white square). (A3) Rate versus notch frequency for actual data (crosses; dashed line), first-order model predictions (unfilled circles; light solid line), and second-order model predictions (filled circles; bold solid line). Q_1 and Q_2 (first- and second-order Q-values, respectively) are shown in the legend. Horizontal shaded bar indicates spontaneous discharge rate. (B1) First-order weighting function for a type IV neuron (BF= 15.9 kHz; same neuron as in Fig. 3B). Data were taken ≈ 10 dB above threshold. (B2) Second-order weighting function for the same type IV neuron. Largest weight values are located off-BF and off-diagonal. (B3) Response rate predictions for the same type IV neuron, plotted as in A3.

Second-order weights for the same chopper and type IV neurons are plotted in Fig. 4A2 and B2, respectively. In these checkerboard plots, on-diagonal values indicate the degree to which a simple quadratic term $S(f_i)^2$ at a single frequency f_i contributes to the overall rate response. Off-diagonal values describe more complex interactions between the indicated frequency pairs. Although the chopper neuron has a single large second-order weight at BF (Fig. 4A2), the type IV neuron has a more variable second-order weight pattern with substantial components off-BF (Fig. 4B2). In both cases, inclusion of these second-order terms in the model (Eq. 1) improves the predicted responses to HRTF stimuli with notches above BF. The overall Q-value rises from 0.46 to 0.76 in the second-order chopper prediction (Fig. 4A3) and from 0.42 to 0.54 in the type IV prediction (Fig. 4B3). These improvements in the model suggest that both systems have significant second-order nonlinearities that are accurately depicted by the weight patterns shown. In the

chopper, the BF component is the most significant and is relatively independent of all other frequency components. The second-order term is simply required to better describe the sigmoidal shape of the input-output function near threshold. In contrast, responses of type IV neurons appear to depend on more complex interactions between frequency components away from BF. For example, the largest second-order weights for this type IV neuron suggest that the response depends on an inhibitory process involving frequency components located 0.25 and 0.75 octaves below BF and an excitatory process at an adjacent frequency. However, the second-order weight estimates somewhat depend on the frequency range over which they are computed, and it is not yet clear how they are to be interpreted.

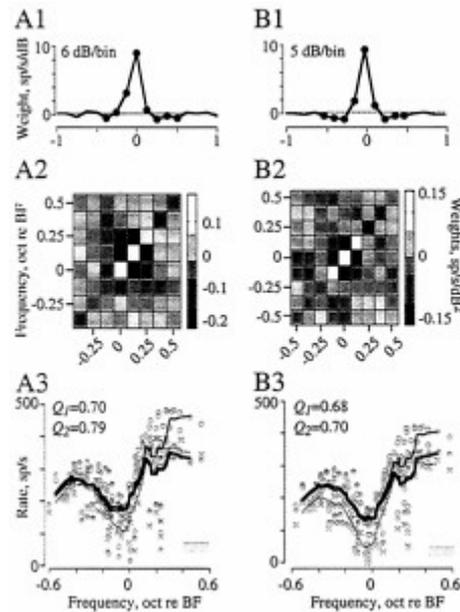


Fig. 5. Weighting functions and notch response predictions for two chopper neurons studied at approximately 40 dB above threshold. The chopper in *A* is the same neuron shown in Figs. 3*A* and 4*A*. The chopper in *B* has a BF of 10.2 kHz. Plot layout is the same as Fig. 4. Note the good match of the predictions to actual data.

In chopper neurons, the noise response properties at levels above threshold are similar to those at levels near threshold. Fig. 5 shows the weighting functions and response predictions for two chopper neurons at ≈ 40 dB above threshold. At this higher level, both neurons exhibit triangular first-order weighting functions that peak at BF. Weaker inhibitory side bands are also present (Fig. 5*A1* and *B1*). Note the large magnitude of the peak weights in the first-order functions. In choppers, peak weights often range between 5 and 10 spikes/s per dB at high levels. Weight values observed in auditory nerve fibers at these levels are substantially smaller (1–2 spikes/s per dB, ref. 33), suggesting that chopper neurons are receiving converging input from multiple auditory nerve fibers. Note also that first-order peak weight values are level-dependent. The neuron shown in Fig. 5*A* is the same as that depicted near threshold in Fig. 4*A*. In general, peak weights are small at low sound levels but monotonically increase in amplitude to a maximum in the middle of the neuron's dynamic range. Saturation effects eventually cause a decrease in peak weight values at higher stimulus levels (not shown).

At levels above threshold, second-order chopper response functions often show significant positive (excitatory) quadratic terms that involve one or a few frequency bins near BF. This typical second-order weight pattern, similar to that observed near threshold, is illustrated in Fig. 5*A2* and *B2*. Note that in both plots, negative second-order weights exist off-BF, which suggests that these particular choppers are weakly inhibited by joint activity of the BF component and an adjacent component above BF.

At levels within the dynamic range of chopper neurons, the transformation of the stimulus spectrum into rate is predominantly a first-order (or linear) transformation of spectral level. Rate minima are produced when the HRTF notch, a bin of minimum spectral energy, is centered at the excitatory peak of the first-order weighting function. Fig. 5*A3* and *B3* shows chopper rate predictions to notch stimuli at moderate levels. Qualitatively, the shape of the rate versus notch frequency relationship is captured by the linear model. Q -values for first-order predictions ($Q_1=0.70$ and 0.68) are high. In both neurons, the linear model does overestimate the actual rate when the notch is located above BF. This discrepancy can be corrected by adding second-order terms to the model; however, the addition of this nonlinearity does not improve the model by much ($Q_2=0.79$ and 0.70).

Spectral Notch Encoding Is a Nonlinear Process in Most Type IV Neurons at Suprathreshold Levels

Unlike chopper response functions at suprathreshold levels, type IV weighting functions are variable in form. In some of the first-order weighting functions, the excitatory (positive) peaks at BF that are observed near threshold are retained with increasing level. In others, these peaks are replaced by wide areas of inhibition. Second-order weighting functions are similarly variable and often exhibit significant off-BF terms. The variability of type IV response functions is apparent in Fig. 6, which shows the weights estimated for two different type IV neurons at 30–40 dB above threshold. Unlike chopper neurons, type IV neurons at suprathreshold levels have first-order functions with large inhibitory regions, relatively small positive peaks, and relatively low overall gains. A wide band of negative (inhibitory) weights below BF, as in Fig. 6*A1*, is a common finding in the weighting functions of type IV neurons. Alternatively, type IV weighting functions may resemble the example in Fig. 6*B1*. This weighting function reveals an excitatory drive at BF and strong inhibition extending above and below BF. Note that the amplitude of first-order weights rarely exceeds a value of 2 in type IV neurons.

The responses of type IV neurons with large inhibitory values in their first-order weighting functions are poorly described by the model in Eq. 1. Actual responses of two type IV neurons to HRTF stimuli are compared with model predictions in Fig. 6*A3* and *B3*. The inhibitory notch responses in the two examples are clear; in both neurons, actual rates fall below spontaneous rate as the notch approaches BF. The first-order model qualitatively fails to predict the depth of notch responses and does not show inhibition. It also yields a positive DC rate offset across all notch frequencies. The second-order model corrects for the DC offset to some extent, but overall predictions remain poor.

Q -values for a first-order model (Q_1) were computed for 19 chopper and 23 type IV neurons across a range of seven levels. Results are compiled in Fig. 7, showing the level dependence of Q_1 for each neuron. The trend in level dependence across each of the two neural populations is summarized by computing the median Q_1 within a 10-dB range centered at the indicated level. At low levels, both chopper and type IV neurons exhibit roughly

equivalent linearity. However, as stimulus level rises, the Q_I continues to rise in choppers but begins to fall in type IV neurons. Choppers often maintain Q_I values above 0.6 across a 30-dB range of high stimulus levels. In contrast, across these same stimulus levels, the Q_I of type IV neurons declines toward 0. This summary corroborates previous studies that have shown that the behavior of type IV neurons is driven by a nonlinear process that emerges 10–20 dB above threshold (24, 29). The results shown in Fig. 6 indicate that this nonlinearity can be described only by a response function of order greater than 2.

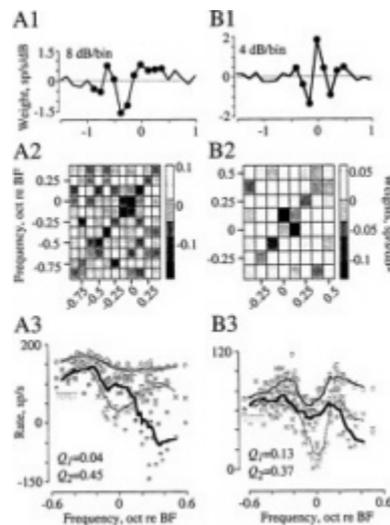


Fig. 6. Weighting functions and notch response predictions for two type IV neurons studied at approximately 30 (A) and 40 (B) dB above threshold. The type IV in A is the same neuron shown in Figs. 3B and 4B. The type IV in B has a BF of 9.9 kHz. Plot layout is the same as Fig. 4. Note the poor match of the predictions to the data.

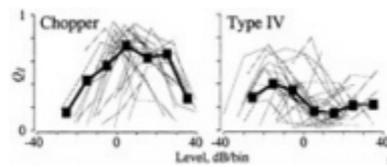


Fig. 7. A summary of Q_I -values obtained in rate predictions of 19 chopper (Left) and 23 type IV (Right) neurons. Q_I -values obtained at multiple levels for a single neuron are connected by a thin gray line. The level dependence is summarized (bold black line) by computing the median Q_I -value within a 10-dB range centered at the indicated level (i.e., abscissa position of the filled square).

Tonotopic Representation Versus Feature Detection

Although responses of chopper and type IV neurons to notches are qualitatively similar, the data in Figs. 5 and 6 show that the mechanisms that govern these responses are quite different. In chopper neurons, the transformation of spectral level into average discharge rate is predominantly a linear (first-order) process that involves a narrow band of frequencies centered on BF. This property suggests that axonal projections from a tonotopic array of chopper neurons transmit, to other auditory areas, a set of rate responses that are homomorphic with spectral shape. In fact, chopper neurons have been shown to produce stable tonotopic representations of vowel spectral shape (8).

Unlike choppers, type IV neurons do not encode spectral information through a homomorphic spectral representation. This is apparent from (i) the complexity and wide bandwidth of their first-order weighting functions, and (ii) the strong nonlinearity in the stimulus-response function. The nature of these nonlinearities has been discussed elsewhere (24, 29) and will not be further elaborated here. It is sufficient to point out that, as Fig. 3 illustrates, type IV neurons give inhibitory responses to both a narrow peak of energy (tone) and a narrow notch of energy located at BF. The rate dependence of type IV neurons to spectral notch position (Figs. 3 and 6), in addition to their wide bandwidth and nonlinear behavior, suggests that these neurons are not simply detectors of spectral level. Type IV neurons provide a second pathway of spectral information transmission—a nonlinear one that signals the presence of a specific complex spectral feature.

The Significance of the System Identification Method

A mathematical representation of a neural system can be a powerful tool for exploring the functional roles of component neurons in a complex acoustic environment. The system identification approach described in this work is a straightforward method for determining the receptive field of an auditory neuron that is responding to noise stimuli. It clarifies the excitatory and inhibitory nature of converging inputs to CN neurons and their location on a tonotopic frequency axis. Moreover, it can be used to study the nature of the information-bearing elements that are processed within the specialized parallel pathways of the auditory system.

The strength of the method lies in the fact that response functions that describe neural receptive fields can be computed simply. The validity of the model weighting functions can be tested by predicting responses to novel stimuli in a manner that is easy and direct. In principle, this method can be modified to study higher-order nonlinearities. Admittedly, however, the latter endeavor would be difficult because it would require the collection of large amounts of stationary data over lengthy time periods.

This method has been shown to work well for VCN chopper neurons. The failure of the system identification model to predict type IV responses is somewhat surprising because the model is essentially a linearization method that should be expected to work over some range of sound level variation. The fact that the model fails at suprathreshold levels suggests that the behavior of type IV neurons is governed by a nonlinear process of high order that cannot be linearized. Although it is possible that such a strong nonlinearity could be linearized by using smaller sound level deviations, such a result would not be particularly interesting because fluctuations of 10–12 dB are typical of the spectral level variations in natural stimuli.

The weighting function method discussed in this paper can serve as a guide to the methods that must be applied in delineating the nature of stimulus representations in a particular neuron. If the behavior of a neuron is found to be linear or nearly linear by using the model in Eq. 1, tonotopic representations sufficiently describe the neuron's spectral

information encoding mechanism. If the behavior cannot be linearized, then receptive fields derived by any single method are unlikely to be informative about the manner in which a neuron encodes arbitrary stimuli. Such neurons must be studied carefully with a variety of stimuli. Considerations of biological function will be important in extracting the important response properties of such neurons.

This work was supported by Grants DC00115 and DC00979 from the National Institute on Deafness and Other Communication Disorders.

1. Merigan, W.H. & Maunsell, J.H.R. (1993) *Annu. Rev. Neurosci.* **16**, 369–402.
2. Rauschecker, J.P. (1997) *Acta Otolaryngol. Suppl.*, **532**, 34–38.
3. Young, E.D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 933–934.
4. Kaas, J.H., Hackett, T.A. & Tramo, M.J. (1999) *Curr. Opin. Neurobiol.* **9**, 164–170.
5. Helfert, R.H., Snead, C.R. & Altschuler, R.A. (1991) in *Neurobiology of Hearing: The Central Auditory System*, eds. Altschuler, R.A., Bobbin, R.P., Clopton, B.M. & Hoffman, D.W. (Raven, New York), pp. 1–25.
6. Young, E.D. (1998) in *The Synaptic Organization of the Brain*, ed. Shepherd, G.M. (Oxford, New York), 4th Ed., pp. 121–157.
7. Peterson, G.E. & Barney, H.L. (1952) *J. Acoust. Soc. Am.* **24**, 175–184.
8. Blackburn, C.C. & Sachs, M.B. (1990) *J. Neurophysiol.* **63**, 1191–1212.
9. Young, E.D., Spirou, G.A., Rice, J.J. & Voigt, H.F. (1992) *Philos. Trans. R. Soc. London B* **336**, 407–413.
10. Middlebrooks, J.C. & Green, D.M. (1991) *Annu. Rev. Psychol.* **42**, 135–159.
11. Huang, A.Y. & May, B.J. (1996) *J. Acoust. Soc. Am.* **100**, 1070–1080.
12. Wightman, F.L. & Kistler, D.J. (1989) *J. Acoust. Soc. Am.* **85**, 858–867.
13. Musicant, A.D., Chan, J.C.K. & Hind, J.E. (1990) *J. Acoust. Soc. Am.* **87**, 757–781.
14. Rice, J.J., May, B.J., Spirou, G.A. & Young, E.D. (1992) *Hear. Res.* **58**, 132–152.
15. Huang, A.Y. & May, B.J. (1996) *J. Acoust. Soc. Am.* **100**, 2341–2348.
16. Irvine, D.R.F. (1986) *The Auditory Brainstem* (Springer, Berlin).
17. Spirou, G.A., May, B.J., Wright, D.D. & Ryugo, D.K. (1993) *J. Comp. Neurol.* **329**, 36–52.
18. Rice, J.J., Young, E.D. & Spirou, G.A. (1995) *J. Acoust. Soc. Am.* **97**, 1764–1776.
19. Spirou, G.A. & Young, E.D. (1991) *J. Neurophysiol.* **65**, 1750–1768.
20. Conley, R.A. & Keilson, S.E. (1995) *J. Acoust. Soc. Am.* **98**, 3223–3234.
21. May, B.J. & Huang, A.Y. (1997) *J. Acoust. Soc. Am.* **101**, 2705–2719.
22. May, B.J., LePrell, G.S. & Sachs, M.B. (1998) *J. Neurophysiol.* **79**, 1755–1767.
23. Berg, B.G. & Green, D.M. (1990) *J. Acoust. Soc. Am.* **88**, 758–766.
24. Nelken, I., Kim, P.J. & Young, E.D. (1997) *J. Neurophysiol.* **78**, 800–811.
25. DiCarlo, J.J., Johnson, K.O. & Hsiao, S.S. (1998) *J. Neurosci.* **18**, 2626–2645.
26. Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. (1986) *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge).
27. Efron, B. & Tibshirani, R.J. (1993) *An Introduction to the Bootstrap* (Chapman & Hall, New York).
28. Nelken, I. & Young, E.D. (1994) *J. Neurophysiol.* **71**, 2446–2462.
29. Nelken, I. & Young, E.D. (1997) *J. Neurophysiol.* **78**, 790–799.
30. Davis, K.A. & Young, E.D. (2000) *J. Neurophysiol.* **83**, 926–940.
31. Voigt, H.F. & Young, E.D. (1990) *J. Neurophysiol.* **64**, 1590–1610.
32. Poon, P.W.F. & Brugge, J.F. (1993) *J. Neurophysiol.* **70**, 655–666.
33. Calhoun, B.M., Miller, R.L., Wong, J.C. & Young, E.D. (1998) in *Psychophysical and Physiological Advances in Hearing*, eds. Palmer, A.R., Rees, A., Summerfield, Q. & Meddis, R. (Whurr, London), pp. 170–177.

CELLULAR MECHANISMS FOR RESOLVING PHASE AMBIGUITY IN THE OWL'S INFERIOR COLLICULUS

José Luis Peña* and Masakazu Konishi

Division of Biology 216-76, California Institute of Technology, Pasadena, CA 91125

Both mammals and birds use the interaural time difference (ITD) for localization of sound in the horizontal plane. They may localize either real or phantom sound sources, when the signal consists of a narrow frequency band. This ambiguity does not occur with broadband signals. A plot of impulse rates or amplitude of excitatory postsynaptic potentials against ITDs (ITD curve) consists of peaks and troughs. In the external nucleus (ICX) of the owl's inferior colliculus, ITD curves show multiple peaks when the signal is narrow-band, such as tones. Of these peaks, one occurs at ITDi, which is independent of frequency, and others at $\text{ITDi} \pm T$, where T is the tonal period. The ITD curve of the same neuron shows a large peak (main peak) at ITDi and no or small peaks (side peaks) at $\text{ITDi} \pm T$, when the signal is broadband. ITD curves for postsynaptic potentials indicate that ICX neurons integrate the results of binaural cross-correlation in different frequency bands. However, the difference between the main and side peaks is small. ICX neurons further enhance this difference in the process of converting membrane potentials to impulse rates. Inhibition also appears to augment the difference between the main and side peaks.

Humans and owls may incorrectly localize a virtual sound source in the hemifield opposite to that of the real source when the signal is narrow-band, such as tones. This problem does not occur when the signal is broadband, such as noise (1, 2). We can explain the cause of this phenomenon by assuming that a process similar to cross-correlation measures the interaural time difference (ITD), which is the cue for sound localization in the horizontal plane (3-5). For example, suppose that the tone in the right ear leads its copy in the left ear by $\frac{1}{4}$ of the period (T). The auditory system delays the neural signals from the right ear by $\frac{1}{4}T$ so that they coincide with those from the left ear, provided that this delay is within the physiological range given by the size of the head. The detection of this coincidence indicates to the brain that the sound is located in the direction encoded by the ITD in the right hemifield. However, a delay of the neural signals from the left ear by $\frac{3}{4}T$ (which equals $\frac{1}{4}T - T$) also causes the two signals to coincide, indicating that the sound source is now in the left hemifield. Thus, narrow band signals may give ambiguous information about sound sources. This problem does not occur with broadband signals, because cross-correlation of them produces a single maximum or peak.

The auditory system does not directly cross-correlate broadband signals but performs this computation on their spectral components. The inner ear breaks down complex signals into their spectral components and encodes the amplitude and phase of each frequency component. Nerve impulses that occur at or near particular phase angles represent phase information. Specific brainstem circuits perform cross-correlation on left and right trains of phase-locked impulses in each frequency band (6-8). Cross-correlation of such signals produces peaks and troughs of impulse rates (e.g., refs. 7, 9, and 10). Connecting the mean discharge rates for different ITDs gives rise to a sinusoidal curve, which will be referred to as the ITD curve (Fig. 1A). It contains one peak at the time lag corresponding to the ITDi (i stands for frequency independent), which represents the source direction, and other peaks at $\text{ITDi} \pm T$, where T is the period of the stimulus tone. This phenomenon is referred to as phaseambiguity and $\text{ITDi} \pm T$ as the phase equivalents (of ITDi). Phase ambiguity occurs irrespective of signal bandwidth, if the neuron is narrowly tuned to frequency, as in brainstem neurons that perform cross-correlation (Fig. 1B). The question here is how the auditory system distinguishes ITDi from its phase equivalents.

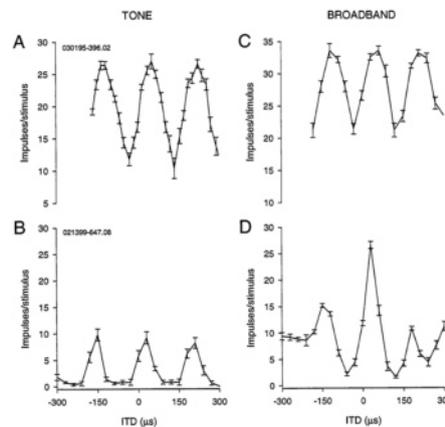


Fig. 1. Phase ambiguity and signal bandwidth. (A) ITD-curve of a nucleus laminaris neuron for a 5.9-kHz tone showing phase ambiguity. (B) The response of the laminaris neuron in A to a broadband signal shows little difference from A. (C) ITD-curve of an ICX neuron for a 5.5-kHz tone. (D) The same neuron as in C stimulated with a broadband signal shows one large main peak and two small side peaks.

If the discrimination between ITDi and its phase equivalents is to take place at the level of single neurons, different frequency bands must converge on single neurons. The owl's auditory system contains such neurons in the external nucleus of the

*To whom reprint requests should be addressed. E-mail: jose@etho.caltech.edu.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19-21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviation footnote: ITD, interaural time difference; ITDi, interaural time difference, frequency independent; ICC, central nucleus of the inferior colliculus; ICX, external nucleus of the inferior colliculus; M-SP, main-side peak; NL, nucleus laminaris; PSP, postsynaptic potential.

inferior colliculus (ICX) (11). The frequency tuning curves of ICX neurons are broad, spanning from 2 kHz to 8 kHz. These neurons show phase-ambiguity in response to ITDs contained in tones, mirroring the behavioral findings (Fig. 1C). Comparison of the ITD curves obtained for different frequencies clearly shows an ITD that is independent of frequency. Furthermore, when the signal is broadband, the same neurons respond to this ITD exclusively or much more strongly than to other ITDs (12) (Fig. 1D). Thus, in the ICX and other higher order neurons, such as those of the optic tectum that receive input from the ICX, the peak at ITDi (“main peak”) is clearly distinguishable from those (“side peaks”) at its phase equivalents, when the signal is broadband (2, 13). In addition, ICX neurons receive inputs from neurons that respond to a common ITD, although they are sensitive to different frequencies. These neurons occur in a tonotopic sequence in the central nucleus of the inferior colliculus and project to the ICX (14). These findings led to a hypothesis that accounts for the elimination of phase ambiguity by frequency convergence (14–16) (Fig. 2). In this model, a single ICX neuron receives inputs from lower order neurons that prefer a common ITD (i.e., ITDi). The phase equivalents of ITDi vary, however, because the neurons are tuned to different frequencies. Addition of the ITD curves across frequency with respect to ITDi gives rise to a tall peak at this ITD and deep troughs on both sides of it. Peaks at the phase equivalents are smaller, because these peaks do not coincide across different frequency bands. When peaks and troughs coincide, the resulting ITD curve shows elevated plateaus flanking the deep troughs on both sides of the main peak. The details of the distribution of these small peaks and the shape of the plateaus should vary according to the frequency preference of the neuron and the spectral composition of the signal. In general, the sum of the results of cross-correlation performed on the spectral components of a signal is equivalent to cross-correlating the whole signal. The efficacy of broadband signals for unambiguous localization of sound by humans has also been explained by across-frequency convergence of ITD information (e.g., refs. 1, 4, and 17–19). ICX neurons appear to use not only additive processes in frequency convergence but also nonlinear mechanisms to discriminate between the main peak and its phase equivalents. These neurons offer unique opportunities to study how these linear and nonlinear operations are carried out at the level of single neurons. We used intracellular methods to see the relationships between the input and output sides of these computations as they occur in these neurons.

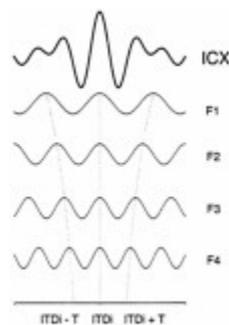


Fig. 2. Model of frequency convergence for elimination of phase ambiguity. ICX neurons sum the results of cross-correlation performed in different frequency bands (F1-F4). Cross-correlation peaks occur at ITDi and at its phase equivalents $ITDi \pm T$, where T is tonal period, which varies with frequency. Integration across frequency gives rise to a tall peak at ITDi. Dotted lines show how the peaks at ITDi line up across frequency, whereas the peaks at $ITDi \pm T$ diverge with frequency.

Materials and Methods

Surgery. Data were obtained by *in vivo* intracellular recording of 59 inferior colliculus space-specific neurons in 14 adult barn owls. These neurons are tuned to combinations of ITD and interaural intensity difference and selective for the direction of sound sources. The protocol for this study followed the National Institutes of Health Guide for the Care and Use of Laboratory Animals and was approved by the Laboratory Animal Care and Use Committee. Owls were anesthetized by i.m. injection of ketamine hydrochloride (25 mg/kg; Ketaset; Phoenix Pharmaceutical, St. Joseph, MO) and diazepam (1.3 mg/kg; Diazepam; Steris, Mentor, OH). An adequate level of anesthesia was maintained with supplemental injections of ketamine. Body temperature was maintained with a heating pad. The skull was immobilized by placing the owl in a stereotaxic head holder such that the palatine ridge was inclined down by 90° with respect to the horizontal bars of the stereotaxic holder. We cemented an inverted T shaped metal plate on the skull to support the head during the experiment without obstructing the ear canals. Before we transferred owls to recovery cages, we injected another dose of diazepam to ensure a smooth recovery from anesthesia, an analgesic (buprenorphine hydrochloride, 0.03 mg, Buprenex; Reckitt & Colman Pharmaceuticals, Richmond, VA), and an antibiotic, (20 mg, Oxytetracycline, Phoenix Pharmaceutical) to protect owls from pains and infection.

Recording sessions began 7–10 days after the implantation of the headplate. Anesthesia was induced and maintained, and antibiotic was given as described above. Ten milliliters of lactated ringer was s.c. injected. All instruments were sterilized. The midbrain was approached through a hole made in the occipital bone through which the bony eminence containing the optic lobe is visible. We made a small hole on this bone to insert electrodes.

Acoustic Stimuli. All experiments were performed in a double-walled sound-attenuating chamber. Acoustic stimuli were delivered by an earphone-assembly consisting of a Knowles ED-1914 receiver as a sound source, a Knowles (Electronics, Franklin Park, IL) BF-1743 damped coupling assembly for smoothing the frequency response of the receiver, and a calibrated Knowles 1939 microphone for monitoring sound pressure levels in the ear canal. The Knowles components were encased in an aluminum cylinder 7 mm in diameter and 8.1 mm in length. The cylinder was inserted into the ear, and the gaps between the earphone assembly and the ear canal were sealed with silicone impression material (Gold Velvet, Earmold and Research Laboratories, Wichita, KS). The calibration data contained the amplitudes and phase angles measures in steps of 100 Hz. Irregularities in the frequency response of each earphone were automatically smoothed by the computer from 2 kHz to 12 kHz.

Acoustic stimuli were digitally synthesized with a Dell Dimension XPS Pro200n computer and delivered by a stereo analog interface (DD1, Tucker-Davis Technologies, Gainesville, FL). ITDs were computed online, whereas interaural level differences were set by two digital attenuators (PA4, Tucker-Davis Technologies), which were controlled by the computer. Tonal stimuli 100 msec in duration, 5 msec linear rise/fall time, were presented once per second. ITD was varied in steps of 30 μ sec. We used the same methods of decorrelating binaural signals that were used previously in this laboratory (16).

Intracellular Recording. Sharp borosilicate glass electrodes filled with 2 M potassium acetate and 4% neurobiotin were used for intracellular recording of space-specific neurons. Analog signals

were amplified (Axoclamp 2A, Axon Instruments, Foster City, CA) and stored in the computer.

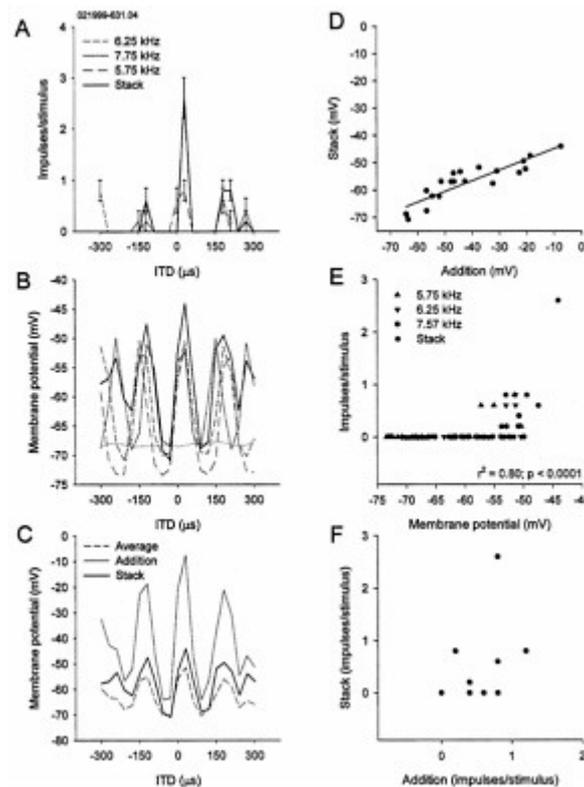


Fig. 3. Frequency integration in an ICX neuron. (A) ITD curves for three different frequencies and one obtained for a combination (stack) of the three. Note the tall main peak and shorter side peaks in the stack curve, (B) Intracellular version of A, plotting evoked PSPs instead of impulse rates as a function of ITD. Note the small differences between the main and side peaks in contrast with A. Resting membrane potential is indicated by dotted line. (C) Neither addition nor average of the three ITD curves predicts the stack curve. (D) The addition of the three curves is correlated with the stack curve in terms of PSPs. (E) Nonlinear relation between impulse rates and PSPs. (F) The absence of correlation between stack and addition in impulse rates.

To measure postsynaptic potentials (PSPs), we used a method of median filtering that eliminates action potentials by replacing each point of the trace by the median of all of the points within a surrounding 5-ms window (20). ITD and frequency tuning and intensity response curves of PSPs were made by custom software (MATLAB 5.3). We calculated the median membrane potential of the first 50 ms of the response and averaged it over five repetitions of the same stimulus. Intracellular best frequency was defined as the frequency that depolarized the cell most at a constant sound intensity 20–30 dB above threshold. Intracellular frequency tuning width was defined as the width of the PSP frequency tuning curve at 30% of the maximum amplitude.

We identified ICX neurons by labeling their axons, which, by definition, project to the optic tectum. The tracer, neurobiotin, was injected by iontophoresis (3 nA-positive, 300-ms current steps, 3 per second for 5 to 30 min). After the experiment, owls were overdosed with Nembutal and perfused with 2% paraformaldehyde. Brain tissue was cut in 60- μ m-thick sections and processed according to standard protocols (21).

Results

The hypothesis of frequency convergence assumes that ICX neurons integrate ITD responses from different frequency bands with respect to ITDi. We examined whether the addition of ITD curves obtained separately for different frequencies predicts the ITD curve for a combination (stack) of the frequencies in 20 neurons. Plotting PSP amplitude against ITD produced a typical ITD curve with peaks and troughs. Fig. 3A and B compares stack curves for impulse rates and PSPs for the same neuron and stimuli. In all neurons, stack curves of PSPs were equal to or larger than either of the ITD curves obtained separately for each frequency. However, the stack curve was always smaller than the addition of the three curves, and larger than their average, as in Fig. 3C. The addition of two or three curves was correlated with the stack curve in most neurons, although the slope of the regression line was small (Fig. 3D). In all cases, the main peak of the stack curve was always taller than the side peaks. Different frequencies appeared to contribute different amounts to the stack curve. The best frequency of a neuron contributed more than other frequencies, although the tuning curves of ICX neurons are broad.

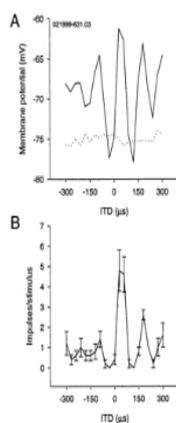


Fig. 4. Small differences between main and side peaks in postsynaptic potentials. (A) ITD curve in PSPs for broadband stimulus. Dotted line shows resting membrane potential. (B) ITD curve in impulse rates for broadband stimulus.

A broadband stimulus induced a large peak at the ITDi and slightly smaller peaks at its phase equivalents. Fig. 4 compares noise-induced ITD curves for PSPs and impulse rates for the same neuron. The presence of large side peaks in PSPs appears to contradict the fact that these neurons are broadly frequency tuned. This phenomenon did not result from any spectral biases in the broadband stimulus used, because the system was calibrated to produce a flat spectral distribution in the ear. Also, the distance between ITD peaks in these neurons was clearly correlated with the neurons' best frequency, which is the frequency that elicits a maximal discharge rate (Fig. 5). The best frequency does not necessarily occupy the midpoint in frequency tuning curve in ICX neurons. We found that the frequency at the midpoint was not as well correlated with the interpeak distance as the best frequency. Despite the apparent discrepancy between the broad frequency tuning and the size of the side peaks in individual neurons, the width of frequency tuning was correlated with the M-SP difference [$100 \times (\text{main peak} - \text{side peak}) \div \text{main peak}$] in the sample of neurons studied; the broader the tuning width, the larger the M-SP difference. The slope of the regression line was larger with the M-SP difference measured with impulse rates than with PSPs (Fig. 6). Also, this correlation existed only when we used the width of frequency tuning measured in PSPs. It may be because tuning widths measured in impulse rates tended to be narrower and more variable (data not shown).

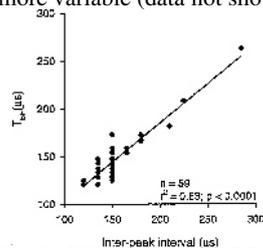


Fig. 5. Correlation between the interpeak interval and the period of best frequency calculated in terms of PSP amplitude (T_{BF}). There is a significant correlation between the two. The solid line represents the linear regression. The number of dots appears smaller than 59 because of overlaps near 150 μsec corresponding to 6.45 kHz.

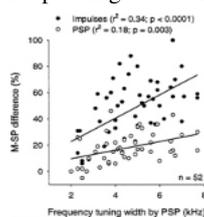


Fig. 6. Difference between main and side peaks (M-SP difference) and filter properties of neurons. Correlation between M-SP difference and frequency-tuning width. M-SP difference for impulse rate (solid circles) and PSP (empty circles) are compared.

Derivation of ITDs by cross-correlation works best if one signal is a time-shifted version of the other signal. Addition of random noise to one or both sides “decorrelates” such a pair of signals. This method can test whether a neuron (or its afferent) behaves as a cross-correlator or not (16). We tested this idea with five neurons for three different degrees of correlation (1, 0.5, and 0), and Fig. 7 shows an example. Decorrelation reduced the PSP amplitude of the main peak and the depths of the troughs. This trend was observed in all cases.

M-SP differences in PSPs are much smaller than those in impulse rates. The question here is how this translation occurs. One mechanism may be to set the threshold of firing between the main and side peaks. Although we saw this in some neurons, there were other neurons in which PSPs for both the main and side peaks exceeded the threshold. Furthermore, the M-SP

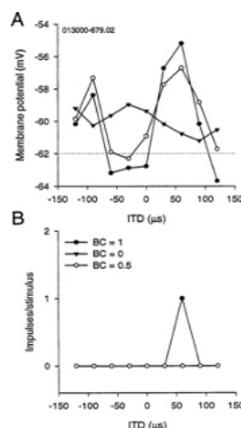


Fig. 7. Effects of decorrelation on ITD-curves. (A) PSP ITD-curves for BC=0, 0.5, and 1. The main peak and one side peak are shown. The amplitude of peaks and troughs decreases with BC. (B) Corresponding impulse-ITD curves to the ones in A. The neuron responds only to the best ITD and BC=1.

differences obtained for PSPs and impulse rates were correlated (Fig. 8). The slope of the regression line was, however, steep, indicating that small differences in PSPs are generally translated into large differences in impulse rates. Although there was a sublinear correlation between the stack curve and the addition of the three curves in PSPs, a highly nonlinear relationship between PSPs and impulse rates was usually observed (cf. Fig. 3E). As a result, the addition of impulse-rate ITD curves for tones was sometimes nonlinearly correlated with the impulse-rate stack curve (cf. Fig. 3F). On the other hand, examples of this nonlinear type were less common than linear ones in which the addition of the two impulse-rate curves could predict the impulse-rate stack curve.

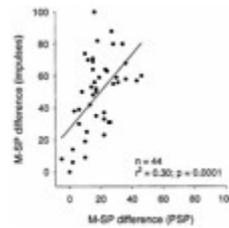


Fig. 8. Relationship between M-SP differences in PSPs and impulse rates. Small changes in M-SP differences in PSPs are associated with large changes in impulse M-SP differences. The solid line represents the linear regression.

In addition to the processes mentioned above, inhibition appears to suppress the side peaks more than the main peak. Fig. 9A shows the effects of increasing sound (noise) levels on impulse rates. Higher levels caused the main peak to increase more than the side peaks. Injection of depolarizing current caused sound-evoked discharge rates to increase. Stimulation with ITDi±T increased this rate less than with ITDi, suggesting that inhibition suppresses the side peaks more than the main peak (Fig. 9B).

Discussion

The site of coincidence detection in owls is the nucleus laminaris (NL) in which neurons are narrowly tuned to frequency (7, 22, 23). The NL projects to the core of the central nucleus of the inferior colliculus directly and indirectly by way of one of the lemniscal nuclei. The NL also projects to a part of the superior olivary nucleus. The results of cross-correlation are conveyed upstream in parallel frequency bands (24). The core projects to the lateral shell of the central nucleus of the contralateral inferior colliculus (25). The convergence of different frequency bands appears to begin in the lateral shell and continues into the adjoining ICX. Both anatomical and physiological lines of evidence show a wide range of frequency convergence in the ICX (14).

ICX neurons integrate ITD-induced PSPs across frequency. The shape of stack curves resembles the addition of cross-correlation performed on its spectral components. This process alone can make the main peak taller than the side peaks, because the PSPs for ITDi coincide across frequency, whereas those for ITDi+T do not. Furthermore, the M-SP difference in impulse rates increases monotonically as a function of signal bandwidth in the ICX and the optic tectum (2, 13). Decorrelation of binaural signals also reduced the main peak in tectal and ICX neurons (ref. 26 and the present work). These findings are consistent with the idea that ICX neurons integrate the results of cross-correlation on separate frequency bands (cf. Fig. 2).

ITD curves of PSPs showed conspicuous periodicity despite the neuron's broadband responses to frequency. However, frequency-tuning curves obtained with tones do not necessarily predict how the neuron responds to broadband signals. The rate at which the side peaks decline with increasing distance from ITDi is a function of the neuron's bandwidth, if the neuron sums the results of cross-correlation in different frequency bands. The broader the width, the steeper the rate. The M-SP difference can represent this rate. The decline is much faster in spike rates than in membrane potential data. The M-SP difference is much larger in impulse rates than in membrane potentials. These changes in the M-SP difference involve nonlinear mechanisms such as inhibition and thresholding. Setting membrane potentials at critical levels for firing by inhibition is such a process. Bicuculline, a γ -aminobutyric acid type A (GABA_A) blocker, raises the discharge level of the entire ITD curve, indicating the possible gain control function of inhibition. Interestingly, the effects of inhibition appear to be greater on the side peaks than on the main peak. Experiments also indicated that the side peaks increased more than the main peak in response to Bicuculline (27, 28). These results are consistent with our interpretation that inhibition reduces side peaks. Taken together, ICX neurons integrate across frequency the outputs of cross-correlation on separate frequency bands and, in addition, perform nonlinear operations on the results of the integration.

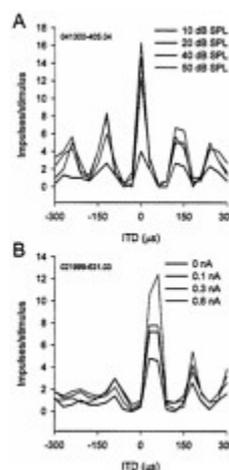


Fig. 9. Differential inhibition on main and side peaks. (A) ITD-curves obtained for different sound levels. (B) ITD-curves during injection of depolarizing current. In both conditions, the main peak increased much more than the side peaks.

Derivation of ITDs by cross-correlation has been extensively studied in the central nucleus (ICC) of the cat. ICC neurons sum the ITD responses in different frequencies within their tuning ranges. The average of these curves can predict the shape of ITD curves for noise stimuli (29, 30). Also, the M-SP difference increase as a function of stimulus bandwidth (29, 31). The effects of decorrelation are also similar between owls and cats (29). In mammals, the medial superior olivary nucleus contains the primary coincidence detectors (9, 10). The question is whether ICC neurons receive input from medial superior olivary nucleus neurons that are tuned to different frequencies. Low frequency primary auditory nerve fibers in cats have a low mean Q_{10} (about 2) (32). This fact suggests that the range of frequency convergence that is observed in the ICC may be taking place in the auditory nerve and relayed to ICC neurons. Moreover, the ITD curves of the cat's medial superior olivary nucleus neurons

obtained for broadband stimuli closely resemble those of ICC neurons in that the main peak is much taller than the side peaks (10). Thus, frequency convergence at the level of the ICC need not be assumed to explain the ITD curves of ICC neurons for noise. In contrast, the ITD responses of the owl's NL neurons to noise do not resemble those of ICX neurons at all (*cf.* Fig. 1), because the latter are much more broadly tuned to frequency.

The broad tuning curves of the owl's ICX neurons may be related to the high frequency range in which the owl extracts both ITDs for azimuth and interaural intensity differences for elevation. ICX neurons selectively respond to combinations of ITD and interaural intensity difference. In this frequency range, the side peaks occur at ITDs that are well within the physiological range of 200 μ sec. In contrast, the side peaks are seldom within the physiological range of ITDs in humans and other mammals that use low frequencies for ITD-based sound localization. The largest M-SP difference (80–100%) achieved by the owl's ICX neurons is perhaps larger than that achieved by the cat's ICC neurons, although quantitative data are not available for the cat. This presumed difference is partly due to the nonlinear processes that ICX neurons perform on the results of frequency integration. This fact does not mean that mammals lack a stage that covers a broader range of frequencies. The ICX of Guinea pigs contains arrays of neurons that are selective for the direction of sound sources (33). Although they require binaural stimuli, their frequency and ITD sensitivity are not known. Study of these neurons with dichotic stimuli would be interesting.

We thank Teresa Nick and Ben Arthur for critically reading the manuscript, Kourosh Saberi for information about human psychoacoustics, Gene Akutagawa for histological assistance, and Chris Malek for computer matters. This work was supported by National Institutes of Health Grant DC00134.

1. Stern, R.M., Zeiberg, A.S. & Trahiotis, C. (1988) *J. Acoust. Soc. Am.* **84**, 156–165.
2. Saberi, K., Takahashi, Y., Farahbod, H. & Konishi, M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6465–6468.
3. Jeffress L.A. (1948) *J. Comp. Physiol. Psych.* **41**, 35–39.
4. Jeffress, L. (1972) in *Foundations of Modern Auditory Theory*, ed. Tobias, J.V. (Academic, New York), Vol II, pp. 349–368.
5. Sayers, B.M. & Cherry, E.C. (1957) *J. Acoust. Soc. Am.* **29**, 973–987.
6. Young, S.R. & Rubel, E.W. (1986) *J. Comp. Neurol.* **254**, 425–459.
7. Carr, C.E. & Konishi, M. (1990) *J. Neurosci.* **10**, 3227–3246.
8. Smith, P.H., Joris, P.X. & Yin, T.C. (1993) *J. Comp. Neurol.* **331**, 245–260.
9. Goldberg, J.M. & Brown, P.B. (1969) *J. Neurophysiol.* **32**, 613–636.
10. Yin, T.C.T. & Chan, C.K. (1990) *J. Neurophysiol.* **64**, 465–487.
11. Knudsen, I.E. & Konishi, M. (1978) *J. Neurophysiol.* **41**, 870–884.
12. Takahashi, T. & Konishi, M. (1986) *J. Neurosci.* **6**, 3413–3422.
13. Mazer, J.A. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 10932–10937.
14. Wagner, H., Takahashi, T. & Konishi, M. (1987) *J. Neurosci.* **7**, 3105–3116.
15. Konishi, M., Takahashi, T., Wagner, H., Sullivan, W.E. & Carr, C.E. (1988) in *Auditory Function: Neurobiological Bases of Hearing*, eds. Edelman, G., Gall, W.E. & Cowan, W.M. (Wiley, New York), pp. 721–745.
16. Albeck, Y. & Konishi, M. (1995) *J. Neurophysiol.* **74**, 1689–1700.
17. Colburn, H.S. (1973) *J. Acoust. Soc. Am.* **54**, 1458–1470.
18. Trahiotis, C. & Stern, R.M. (1994) *J. Acoust. Soc. Am.* **96**, 3804–3806.
19. Shackleton, T.M., Meddis, R. & Hewitt, M.J. (1992) *J. Acoust. Soc. Am.* **91**, 2276–2279.
20. Jagadeesh, B., Wheat, H.S., Kontsevich, L.L., Tyler, C.W. & Ferster, D. (1997) *J. Neurophysiol.* **78**, 2772–2789.
21. Kita, H. & Armstrong, W. (1991) *J. Neurosci. Methods* **37**, 141–150.
22. Pena, J.L., Viète, S.M., Albeck, Y. & Konishi, M. (1997) *J. Neurosci.* **17**, 1815–1824.
23. Viète, S.M., Peña, J.L. & Konishi, M. (1997) *J. Neurosci.* **17**, 1815–1824.
24. Takahashi, T. & Konishi, M. (1988) *J. Comp. Neurol.* **274**, 190–211.
25. Takahashi, T., Wagner, H. & Konishi, M. (1989) *J. Comp. Neurol.* **281**, 545–554.
26. Saberi, K., Takahashi, Y., Konishi, M., Albeck, Y., Arthur, B.J. & Farahbod, H. (1998) *Neuron* **21**, 789–798.
27. Fujita, I. & Konishi, M. (1991) *J. Neurosci.* **11**, 722–739.
28. Mori, K. (1997) *Hearing Res.* **111**, 22–30.
29. Yin, T.C.T. & Chan, C.K. (1986) in *Auditory Function: Neurobiological Bases of Hearing*, eds. Edelman, G., Gall, W.E. & Cowan, W.M. (Wiley, New York), pp. 385–430.
30. Yin, T.C.T., Chan, C.K. & Irvine, D.R. F. (1986) *J. Neurophysiol.* **55**, 280–300.
31. Chan, J.C.K., Yin, T.C.T. & Musicant, A.D. (1987) *J. Neurophysiol.* **58**, 543–561.
32. Kiang, N.Y.-S. (1965) *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA).
33. Binns, K.E., Grant, S., Whittington, D.J. & Keating, M.J. (1992) *Brain Res.* **589**, 231–242.

SUBDIVISIONS OF AUDITORY CORTEX AND PROCESSING STREAMS IN PRIMATES

Jon H. Kaas*[†] and Troy A. Hackett[‡]

Departments of [†]Psychology and [‡]Hearing and Speech Sciences, Vanderbilt University, Nashville, TN 37240

The auditory system of monkeys includes a large number of interconnected subcortical nuclei and cortical areas. At subcortical levels, the structural components of the auditory system of monkeys resemble those of nonprimates, but the organization at cortical levels is different. In monkeys, the ventral nucleus of the medial geniculate complex projects in parallel to a core of three primary-like auditory areas, AI, R, and RT, constituting the first stage of cortical processing. These areas interconnect and project to the homotopic and other locations in the opposite cerebral hemisphere and to a surrounding array of eight proposed belt areas as a second stage of cortical processing. The belt areas in turn project in overlapping patterns to a lateral parabelt region with at least rostral and caudal subdivisions as a third stage of cortical processing. The divisions of the parabelt distribute to adjoining auditory and multimodal regions of the temporal lobe and to four functionally distinct regions of the frontal lobe. Histochemically, chimpanzees and humans have an auditory core that closely resembles that of monkeys. The challenge for future researchers is to understand how this complex system in monkeys analyzes and utilizes auditory information.

The auditory system of mammals includes a large number of interconnected nuclei and cortical areas (Fig. 1). The complexity of this array of interacting structures presents a challenge to researchers to determine how this system analyzes and uses auditory information. Although most of the component subcortical nuclei can be identified in a range of studied mammals, considerable variability in at least cortical organization appears to exist in different lines of evolution, and this variability amplifies the task from understanding the auditory system to understanding auditory systems and their variability.

We have been studying the auditory system of primates, partly because we are interested in how our auditory system works, but also because primates vary in size, auditory behavior, and brain complexity. Thus, we wonder what components of the auditory system are basic and similar across the major taxonomic groups of primates, and what features vary. In particular, we are concerned about auditory cortex, because different lines of evolution vary considerably in cortical organization. For example, whereas both cats and monkeys have quite a number of visual areas, perhaps 30 or more in monkeys, they share only a few (see ref. 2). Most visual areas have emerged independently as both carnivores and primates evolved more complex visual systems. Although the organization of visual cortex in primates has been intensively studied in many laboratories over the last 30 years, efforts to determine the subdivisions and interconnections of auditory cortex have been much more limited. Yet considerable progress has been made, especially recently. Here we review theories of the organization of auditory cortex in primates. These theories have been based largely on studies of the tonotopy, connections, and architecture of auditory cortex in monkeys, but additional relevant information is starting to emerge from investigations of cortical organization in prosimian galagos, histochemical studies in chimpanzees and humans, and noninvasive functional studies in humans.

The Core Areas of Auditory Cortex

Originally, auditory cortex of monkeys was thought to be organized much as in cats, with a single primary area, AI, in the cortex of the lower bank of the lateral sulcus and a second area, All, deeper in the sulcus (e.g., ref. 3). This concept fits well with the early view that auditory, somatosensory, and visual systems all have two fields. However, we now know that primates have a number of sensory representations for each modality, and several somatosensory and auditory fields can be considered primary or primary like in character. In the auditory system, three fields have similar primary-like features. These fields constitute the auditory core, which is surrounded by a belt of secondary fields, and a more lateral parabelt of fields at a third level of processing (Fig. 2). The core includes a most caudal AI, a more rostral field R, and an even more rostral temporal field, RT (Fig. 3). These fields are distinguished from each other by having different systematic representations of the cochlea (6–12). In caudal AI, neurons are best activated by high-frequency tones, whereas neurons in rostral AI are best activated by low-frequency tones. The lines of isorepresentation along rows of neurons across AI having similar best frequencies are curved so that neurons deeper in the fissure have higher best frequencies. The tonotopic organization of R is reversed from that in AI, so that low frequencies are represented caudally in R, and higher frequencies are represented rostrally. Again, the lines of isorepresentation are curved so that neurons deeper in the lateral sulcus have higher best frequencies. The tonotopic organization of RT has not been studied adequately, but it appears that high frequencies are represented caudally and low frequencies, rostrally.

Each of the core areas has features that are characteristic of primary sensory cortex. First, neurons in all three areas respond well and with short latencies to pure tones. The neurons respond best to specific (best) frequencies, and they have narrow frequency-response curves. Second, all three areas receive dense thalamic inputs from the principal or ventral nucleus of the medial geniculate complex (e.g., refs. 9–11, 13–19). These inputs appear to be in parallel so that the ablation of one of the fields does not deactivate the others (20). Third, all three fields have the architectonic characteristics of primary sensory cortex (5, 11, 21, 22). Thus, the fields have a well-developed layer 4 of granule cells and a dense band of myelinated fibers in the middle layers of cortex. In addition, the middle layers of the core densely express the metabolic enzyme, cytochrome oxidase (CO), the

*To whom reprint requests should be addressed at: Department of Psychology, Vanderbilt University, 301 Wilson Hall, 111 21st Avenue South, Nashville, TN 37240. E-mail: Jon.Kaas@vanderbilt.edu.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

enzyme for deactivating the neurotransmitter/neuromodulator acetylcholine (acetylcholinesterase, AChE), and the calcium-binding protein, parvalbumin (Pb) (Fig. 4). Primary sensory areas characteristically express large amounts of CO, AChE, and Pb (e.g., refs. 5, 23–25), although AChE may be more obvious in developing primary areas in some mammals (26, 27). The persistence of large amounts of AChE in the auditory core of adult monkeys is interesting in that acetylcholine is a neurotransmitter associated with developmental plasticity (e.g., ref. 28). Possibly, primary auditory cortex of adult monkeys is especially plastic, so that neurons are capable of changing their

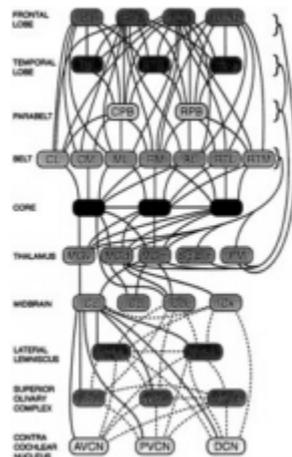


Fig. 1. Cortical and subcortical connections of the primate auditory system. Major cortical and subcortical regions are color coded. Subdivisions within a region have the same color. Solid black lines denote established connections. Dashed lines indicate proposed connections based on findings in other mammals. Joined lines ending in brackets denotes connections with all fields in that region. The belt region may include an additional field, MM (see Fig. 5). Abbreviations of subcortical nuclei: AVCN, anteroventral cochlear nucleus; PVCN, posteroventral cochlear nucleus; DCN, dorsal cochlear nucleus; LSO, lateral superior olivary nucleus; MSO, medial superior olivary nucleus; MNTB, medial nucleus of the trapezoid body; DNLL, dorsal nucleus of the lateral lemniscus; VNLL, ventral nucleus of the lateral lemniscus; ICc, central nucleus of the inferior colliculus; ICp, pericentral nucleus of the inferior colliculus; ICdc, dorsal cortex of the inferior colliculus; ICx, external nucleus of the inferior colliculus; MGv, ventral nucleus of the medial geniculate complex; MGd, dorsal nucleus of the medial geniculate complex; MGm, medial/magnocellular nucleus of the medial geniculate complex; Sg, supragenicular nucleus; Lim, limitans nucleus; PM, medial pulvinar nucleus. Abbreviations of cortical areas: AI, auditory area I; R, rostral area; RT, rostrotemporal area; CL, caudolateral area; CM, caudomedial area; ML, middle lateral area; RM, rostromedial area; AL, anterolateral area; RTL, lateral rostrotemporal area; RTM, medial rostrotemporal area; CPB, caudal parabelt; RPB, rostral parabelt; Tpt, temporoparietal area; TS_{1,2}, superior temporal areas 1 and 2. Frontal lobe areas numbered after the tradition of Brodmann and based on the parcellation of Preuss and Goldman-Rakic (1): 8a, periarculate; 46d, dorsal principal sulcus; 12vl, ventrolateral area; 10, frontal pole; orb, orbitofrontal areas.

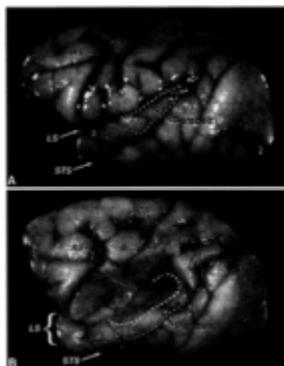


Fig. 2. Lateral view of the macaque cerebral cortex. (A) The approximate location of the parabelt region on the superior temporal gyrus (dashed orange line). (B) Dorsolateral view of the same brain as in A after removal of the overlying parietal cortex, exposing the ventral bank of the lateral sulcus and insula. The approximate locations of the core region (solid red line), caudal and lateral portions of the belt region (dashed yellow line), and the parabelt region (dashed orange line) are shown. The medial portion of the belt region and the rostral pole of the core in the ventral circular sulcus are not visible. Dashed black line defines portion of cortex cut away. AS, arcuate sulcus; CS, central sulcus; INS, insula; LS, lateral sulcus; STG superior temporal gyrus; STS, superior temporal sulcus. Adapted from ref. 4.

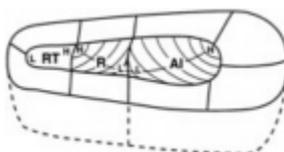


Fig. 3. Tonotopic organization in the core. Auditory core fields (AI, R, RT) are surrounded by belt fields (not labeled). Curved lines within each field depict isofrequency contours. High- (H) frequency acoustic stimuli are represented caudomedially in AI, rostromedially in R. Low- (L) frequency stimuli are represented rostrally in AI, caudolaterally in R. Tonotopic organization in RT is not as certain but may mirror that found in R. See Fig. 1 for abbreviations.

response characteristics (29–31). Although the three core areas are similar in architecture and responsiveness to tones, they are unlikely to be identical in how they process auditory information. Indeed, a number of differences in response characteristics between neurons in AI and R have already been reported (32). In addition, the architectonic features of AI and R are quite similar, but they are somewhat muted in RT. Thus, RT is the least certain member of the core.

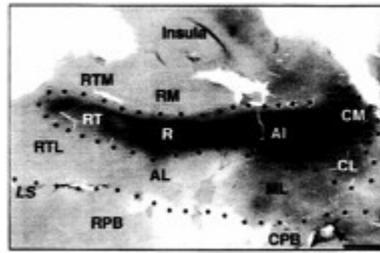


Fig. 4. Architectonic fields in auditory cortex. Macaque brain section flattened and cut parallel to pial surface at 40 μm . Parvalbumin immunohistochemistry. The core fields are the most darkly stained. The caudal belt fields (ML, CL, CM) are moderately dark. Scale bar=5 mm. Adapted from ref. 5. See Fig. 1 for abbreviations.

The cortical connections of the core areas (see refs. 10, 11, 21, 22, 33) are somewhat unusual for primary sensory cortex. Each core area densely interconnects with its neighbor, and even AI and RT have some interconnections (Fig. 5). Thus, core areas must influence each other strongly. In addition, each core area projects to a collection of adjacent belt areas, and the core areas appear to be responsible for activating those belt areas (20, 32). The ipsilateral cortical connections of the core are exclusively, or nearly so, with the narrow surrounding belt of cortex. There are few or no long projections to more distant auditory fields. This means that the belt is an obligatory second stage of cortical processing. The belt is not bypassed, and more distant fields (e.g., parabelt) do not have direct access to core information. In contrast, primary visual cortex, V1, projects both to a surrounding belt, V2, and to a number of more distant visual areas, especially quite distant middle temporal visual area (see ref. 34). Area 3b of somatosensory cortex projects to bordering areas 3a and 1, but also more distantly to areas 2, S2, PV, and even motor cortex (e.g., refs. 35, 36). Thus, primary sensory information is more widely distributed in the somatosensory and visual systems, whereas the auditory system clearly separates three levels of cortical processing.

The connection pattern of the auditory core is unusual in another way. The core areas also project callosally to the core of the other cerebral hemisphere, where the most dense terminations appear to be in tonotopically matched locations of the same area, and to the adjacent belt. Unlike primary visual cortex (V1) and primary area 3b of somatosensory cortex of monkeys, which have large regions of few or no callosal connections, the auditory core has major interhemispheric connections.

The clear manner in which chemoarchitectonic preparations distinguish the core from the belt in monkeys allowed us to examine other primates for the existence of a core. The core in brain sections from temporal cortex of humans and chimpanzees is quite distinct (unpublished observations; refs. 37, 38) and can be identified and delineated with great precision (Fig. 6). In both chimpanzees and humans, the areal extent of the core is greater than in macaque monkeys, but the core in all three primates has the same elongated shape. This suggests that chimpanzees and humans also have three areas within the core (AI, R, RT), but direct evidence is lacking. We do know from studies in humans that a tonotopic pattern of organization can be revealed crudely in the region of the core. A variety of studies using evoked potentials (39, 40), magnetoencephalography (41–52), positron emission tomography (53–56), and functional magnetic resonance imaging (57–60) have produced convincing evidence of tonotopic organization in the human transverse temporal gyrus of Heschl (TTG). As observed in monkeys, high frequencies are represented in the posteromedial TTG of humans, whereas lower frequencies generate activity in anterolateral TTG. Further, apparent reversals in the tonotopic gradient suggest that more than one cochleotopic field may exist in the TTG (45), consistent with evidence for multiple tonotopic fields in the core region of primates.

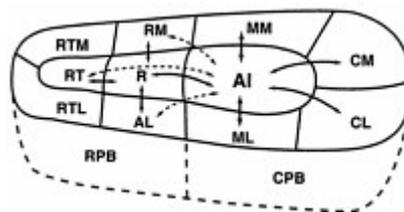


Fig. 5. Auditory cortical connections of AI. Area AI, as well as other core areas, has dense reciprocal connections with adjacent areas of the core and belt (solid lines with arrows). Connections with nonadjacent fields are less dense (dashed lines with arrows). The core has few, if any, connections with the parabelt or more distant cortex. See Fig. 1 for abbreviations.

The Auditory Belt Cortex

The auditory belt is the narrow 2- to 3-mm fringe of cortex immediately surrounding the core with dense interconnections with the core. The belt appears to receive only sparse inputs from the ventral nucleus of the medial geniculate complex, with most of its thalamic inputs coming from the dorsal and medial (magnocellular) divisions of the complex. Despite these thalamic inputs, the belt seems to depend on core inputs for activation. However, this dependence has been demonstrated directly only in one subdivision of the belt, the caudomedial area (CM). Ablation of the core abolishes responses to auditory stimuli in CM (20).

The belt appears to consist of about eight auditory areas, each with a distinct representation of the cochlea. The evidence for these separate areas is largely of two types. First, the core areas connect most densely with adjacent portions of the belt (Fig. 7). This means that the most dense projections to the caudal fields of the belt are from AI, whereas more rostral belt fields get their densest inputs from R or RT. Furthermore, portions of the belt get their densest inputs from immediately adjoining portions of the core, suggesting input patterns that would support multiple, but crude, representations of the cochlea. More specifically, the connection patterns support the possibility of three to four medial belt fields and four lateral belt fields with different sequences of tonotopic organization. These suppositions are supported by the results of microelectrode recordings from the belt. Although neurons in the belt respond much less vigorously to tones than neurons in the core, they respond well enough to indicate that tonotopic gradients in belt areas parallel those of adjacent core areas (12, 61). Neurons in the lateral belt respond better to narrow bands of noise than pure tones, and the center frequency of the band can be varied to indicate best frequencies for neurons (61). These studies provide evidence for three tonotopic representations lateral to the core, the caudolateral area (CL), medio

lateral area (ML), and the anterior lateral area (AL). However, microelectrode recording studies have provided only limited evidence for a tonotopic representation in CM (12, 20, 32), and other medial belt areas have not been studied adequately. Thus, these divisions of the belt are more tentatively proposed and are largely based on connection patterns (see ref. 5). Although the belt is architecturally quite distant from the core, differences between proposed belt areas are not so obvious. CL and CM are somewhat darker than other belt areas in brain sections processed for parvalbumin.

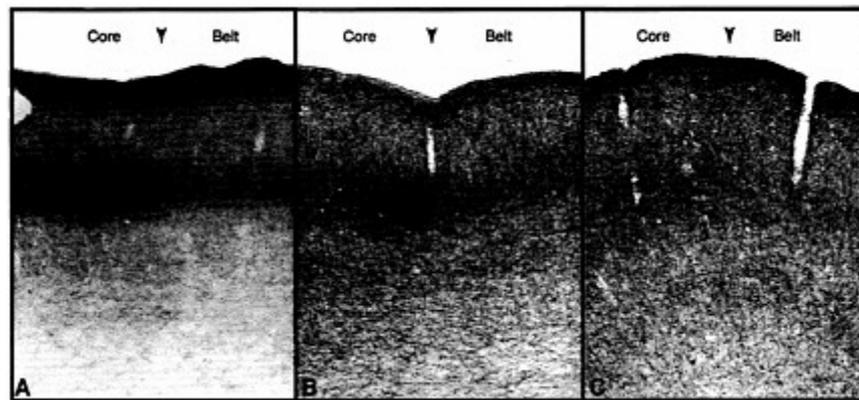


Fig. 6. Architectonic identification of core and belt regions. Coronal sections showing the border between the core and belt regions of auditory cortex (arrowheads). Acetylcholinesterase histochemistry. (A) Macaque monkey; (B) chimpanzee; (C) human. Compared with the belt region, the density of acetylcholinesterase in the middle cortical layers (IIIc and IV) is particularly high in the core.

Besides connections with the adjoining core, belt areas connect with adjoining belt areas, more distant belt areas, the parabelt region, and to the frontal lobe (Fig. 7). Only the connections of the more accessible lateral belt areas have been studied directly, so all of the targets of the medial belt are not known. Nevertheless, it is clear that belt areas are widely interconnected with each other, and they distribute principally to the parabelt, a third distinct level of cortical processing.

The Lateral Parabelt of Auditory Cortex

The parabelt of auditory cortex, just lateral to the lateral belt (Fig. 2), is defined as that region of the temporal lobe where injections of tracers label large numbers of neurons in the belt but few or no neurons in the core (5). The full extent of the parabelt has not been determined accurately, but injections rostral to the parabelt in the $TS_{1,2}$ region and caudal to the parabelt in the Tpt region label the parabelt and other regions of cortex but not the belt. The parabelt may have functionally distinct subdivisions, but little is now known about how to divide the parabelt. Subdivisions are not obvious in the architecture, and systematic recordings with microelectrodes have not yet been attempted. We have tentatively divided the parabelt into rostral and caudal halves on the basis of differences in connections (Fig. 8). The rostral parabelt (RPB) connects largely with the rostral belt areas, whereas the caudal parabelt (CPB) largely connects with the caudal belt areas. However, both divisions connect with the rostromedial area (RM) of the belt. Callosal connections are with largely homotopic portions of the contralateral parabelt and roughly matching levels of the medial and lateral belt. Although the parabelt gets some inputs from the dorsal and medial divisions of the medial geniculate complex, it gets other thalamic inputs from the supragenulate nucleus, nucleus limitans, and medial pulvinar. The parabelt neurons likely depend on belt inputs rather than thalamic inputs for suprathreshold auditory activation.

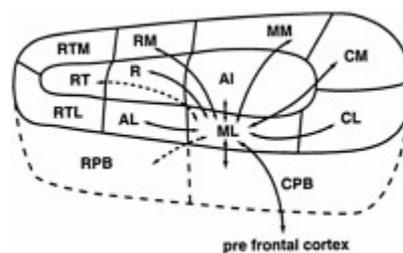


Fig. 7. Auditory cortical connections of ML. Area ML, and other belt areas, have dense connections with adjacent areas of the core, belt, and parabelt (solid lines with arrows). Connections with nonadjacent fields tend to be less dense (dashed lines with arrows). The belt fields also have topographically organized connections with functionally distinct fields in prefrontal cortex. Abbreviations defined in Fig. 1.

Targets of the Parabelt: Additional Levels of Auditory Processing

The parabelt is interconnected with adjacent portions of the temporal and parietal lobe and with several regions of the frontal lobe. These target regions can be considered components of an additional fourth level of auditory processing, with thalamic inputs from the supragenulate, limitans, and medial pulvinar nuclei (Fig. 1; refs. 62–64).

Much of the parabelt connects with nearby cortex of the upper and lower banks of the superior temporal sulcus. The caudal end of this sulcus is occupied by visual areas, such as the middle temporal visual area, and there is no evidence for direct parabelt connections with these visual areas. However, more rostral cortex in the superior temporal sulcus has been referred to as the superior temporal polysensory cortex (STP; see ref. 65 for review) where neurons respond to auditory, visual, and even somatosensory stimulation. The functions of this polysensory cortex are unknown, but visual stimuli are known to influence

strongly the perceived locations of sound sources, and bimodal auditory and visual neurons in STP would be well suited for this interaction. The rostral parabelt connects with more rostral portions of the superior temporal gyrus. On the basis of connections (5), the latter is likely to be largely auditory in function. The caudal parabelt connects with cortex of the caudal end of the superior temporal gyrus, which also may be largely auditory in function, although a nearby region of posterior parietal cortex, 7a, has somatosensory and multimodal functions and neurons involved in reaching for objects (66, 67). Area 7a projects to premotor areas of the frontal lobe that are also involved in guiding reach.

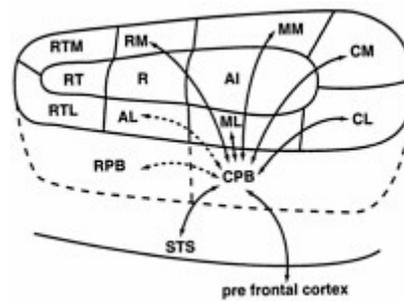


Fig. 8. Auditory cortical connections of CPB. Parabelt area CPB, as well as RPB, has dense connections with adjacent areas of the belt and RM in the medial belt (solid lines with arrows). Connections with nonadjacent fields of the belt tend to be less dense (dashed lines with arrows). The parabelt fields have few, if any, connections with the core areas. The parabelt also has connections with the polysensory areas in the superior temporal sulcus (STS) and with functionally distinct fields in prefrontal cortex. Abbreviations defined in Fig. 1.

The parabelt also projects to four major regions of the frontal lobe (Fig. 9; refs. 4, 69, 70). These include cortex near or within the frontal eye field. This region of cortex (area 8a) is important for directing gaze toward objects of interest. Clearly, sounds from locations in space often would be of visual interest. Other projections are to dorsolateral prefrontal cortex of the principal sulcus in monkeys (area 46d). This cortex is thought to be important in the short-term temporary storage of information, sometimes called working memory (71, 72). Neurons in some locations in this prefrontal region respond to auditory stimuli (73–76). The third projection zone in the frontal lobe is more rostral and ventral (area 12v1). This region of cortex may be involved in the multimodal characterization of objects. In the ventrolateral prefrontal cortex, a number of single-unit studies indicate that this region subserves working memory for nonspatial tasks, such as stimulus recognition (see ref. 4). The fourth target of the parabelt is in orbital-frontal cortex, which is a multimodal region with a role in assigning value to stimuli. The region is associated with the reward system, and it is considered emotive or motivational in function (77–81). Many neurons in orbital-frontal cortex respond to auditory stimuli (82).

On the basis of specific patterns of connections between temporal and frontal cortex, some investigators have distinguished separate pathways for processing spatial and nonspatial auditory information (e.g., refs. 69, 70). These data suggest that spatial (i.e., areas 8a, caudal 46d) and nonspatial (i.e., areas 10, 12v1, rostral 46d) domains in prefrontal cortex are the targets of separate processing streams originating in caudal and rostral fields of nonprimary auditory cortex, respectively. The connection patterns support this distinction, in general, but raise the possibilities of additional streams and significant interactions between streams (83).

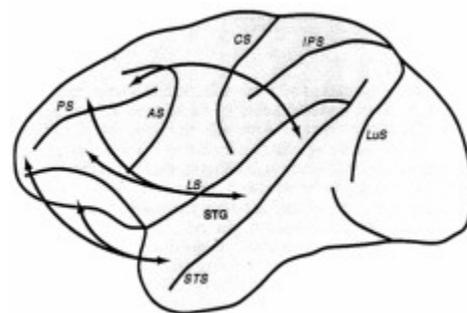


Fig. 9. Prefrontal connections of auditory cortex in macaque monkeys. Arrows summarize the topographic connections of the lateral belt and parabelt auditory regions with functionally distinct areas of prefrontal cortex. The targets of caudal auditory fields favor spatial domains of prefrontal cortex (e.g., 8a, caudal 46d), whereas more rostral fields exhibit stronger connections with nonspatial regions (e.g., areas 12v1, 10, mediofrontal, orbitofrontal). Connections with intermediate temporal fields tend to overlap more. See Fig. 1 and ref. 83 for more detailed connections. Adapted from ref. 68, with permission from S.Karger A, Basel.

Evidence for Levels of Cortical Processing in Humans

In addition to evidence for a core of distinct tonotopic fields in humans, there is also evidence in this literature for hierarchical processing in auditory cortex such that auditory-related activity in cortical fields outside of the core region can be dissociated from activity within by using various techniques (84–105). Howard *et al.* (104) recorded evoked potentials from auditory cortex of patients undergoing surgery for intractable epilepsy. They were able to dissociate auditory responsive fields in Heschl's gyrus (HG) and the posterolateral superior temporal gyrus (PLST) on the basis of differences in responses to auditory stimuli and sensitivity to anesthesia. In addition, short-latency potentials evoked by electrical stimulation of HG were recorded from PLST, suggesting activation via direct or indirect connections between these areas. Scheich *et al.* (96) used a foreground-background decomposition task to dissociate three distinct auditory cortical regions with functional MRI (fMRI). The activated regions on the superior temporal gyrus (T1, T2, and T3) corresponded to distinct cytoarchitectonic fields (KA, PaAi, and PaAe, respectively) of human auditory cortex (37), which are comparable in relative position to core, lateral belt, and parabelt fields in monkeys. These results compare well with those of Binder *et al.* (105), in which distinct foci of peak activation were resolved in fMRI studies with speech and nonspeech stimuli. All auditory stimuli produced equivalent activations in the medial half of HG, but activation in the surrounding fields varied. White noise activation was centered on the HG with some spread into the surrounding fields, whereas activation produced by frequency-modulated tones was more expansive laterally than for noise. The extent of activation was greatest for speech stimuli, irrespective of linguistic significance, spreading ventrolaterally into cortex of the superior temporal sulcus. The authors concluded that such contrasts provided support for a hierarchical model of auditory word processing in the human temporal lobe based on increasing complexity and integration of temporal and spectral features. Overall, the results of functional studies of auditory cortex in humans indicate the presence of multiple hierarchically

arranged fields, consistent with key features of the primate model.

Conclusions

The auditory systems of all mammals include a large number of complexly interconnected auditory nuclei, and at least several areas of auditory cortex. Although the subcortical auditory system of primates and other animals may have similar component nuclei, further study will likely reveal specializations in types and numbers of neurons and connections. Thus, we do not yet know how similar or different the subcortical auditory systems of various animals might be. At the cortical level, primates appear to have evolved an elaborate network of areas that is quite different from the networks in other mammals. Although most of our present understanding of auditory cortex in primates depends on studies in monkeys, there is some evidence that early levels of cortical processing are similar in monkeys, chimpanzees, and humans, although the human auditory system undoubtedly includes specializations for language. Our growing understanding of the complexity of auditory cortex in monkeys opens the door for many productive studies of the nature of processing within the network.

The auditory cortex of monkeys includes a network of at least 20 interconnected auditory areas or multimodal regions. The cortical system is hierarchical with at least four distinct levels of processing, with several areas at each level. Information is widely distributed within levels, between areas of the same level of opposite cerebral hemispheres, and between areas of one level and the next. Yet, there is enough separation in distributions of connections to suggest the existence of at least distinct rostral and caudal processing streams. The existence of multiple areas at even the primary level of cortical processing has obvious implications. Because areas of the same level have similar connections, they may have overlapping functions. Thus, the system should be able to compensate largely for losses produced by lesions that leave some of the areas of any level intact, but the highly serial nature of processing from level to level indicates that extensive damage to any level would be devastating.

1. Preuss, T.M. & Goldman-Rakic, P.S. (1991) *J. Comp. Neurol.* **310**, 429–474.
2. Kaas, J.H. & Krubitzer, L.A. (1991) in *Neuroanatomy of Visual Pathways and Their Retinotopic Organization*, eds. Dreher, B. & Robinson, S.R. (MacMillan, London), pp. 302–359.
3. Walzl, E.M. (1947) *Laryngoscope* **57**, 778–787.
4. Hackett, T.A., Stepniewska, I. & Kaas, J.H. (1999) *Brain Res.* **817**, 45–58.
5. Hackett, T.A., Stepniewska, I. & Kaas, J.H. (1998) *J. Comp. Neurol.* **394**, 475–495.
6. Merzenich, M.M. & Brugge, J.F. (1973) *Brain Res.* **50**, 275–296.
7. Imig, T.J., Ruggero, M.A., Kitzes, L.F., Javel, E. & Brugge, J.F. (1977) *J. Comp. Neurol.* **171**, 111–128.
8. Aitkin, L.M., Merzenich, M.M., Irvine, D.R.F., Clarey, J.C. & Nelson, J.E. (1986) *J. Comp. Neurol.* **252**, 175–185.
9. Leuthke, L.E., Krubitzer, L.A. & Kaas, J.H. (1989) *J. Comp. Neurol.* **285**, 487–513.
10. Morel, A. & Kaas, J.H. (1992) *J. Comp. Neurol.* **318**, 27–63.
11. Morel, A., Garraghty, P.E. & Kaas, J.H. (1993) *J. Comp. Neurol.* **335**, 437–459.
12. Kosaki, H., Hashikawa, T., He, J. & Jones, E.G. (1997) *J. Comp. Neurol.* **386**, 304–316.
13. Mesulam, M.M. & Pandya, D.N. (1973) *Brain Res.* **60**, 315–333.
14. Burton, H. & Jones, E.G. (1976) *J. Comp. Neurol.* **168**, 249–302.
15. Fitzpatrick, K.A. & Imig, T.J. (1978) *J. Comp. Neurol.* **177**, 537–556.
16. Aitkin, L. M., Kudo, M. & Irvine, D.R.F. (1988) *J. Comp. Neurol.* **269**, 235–248.
17. Pandya, D.N., Rosene, D.L. & Doolittle, A.M. (1994) *J. Comp. Neurol.* **345**, 447–471.
18. Hashikawa, T., Rausell, E., Molinari, M. & Jones, E.G. (1995) *Brain Res.* **544**, 335–341.
19. Molinari, M., Dell'Anna, M.E., Rausell, E., Leggio, M.G., Hashikawa, T. & Jones, E.G. (1995) *J. Comp. Neurol.* **362**, 171–194.
20. Rauschecker, J.P., Tian, B., Pons, T. & Mishkin, M. (1997) *J. Comp. Neurol.* **382**, 89–103.
21. Galaburda, A.M. & Pandya, D.N. (1983) *J. Comp. Neurol.* **221**, 169–184.
22. Jones, E.G., Dell'Anna, M.E., Molinari, M., Rausell, E. & Hashikawa, T. (1995) *J. Comp. Neurol.* **362**, 153–170.
23. Tootell, R.B.H., Hamilton, S.L. & Silverman, M.S. (1985) *J. Neurosci.* **5**, 2786–2800.
24. Fitzpatrick, D. & Diamond, I.T. (1980) *J. Comp. Neurol.* **194**, 703–719.
25. Blumcke, I., Hof, P.R., Morrison, J.H. & Celio, M.R. (1990) *J. Comp. Neurol.* **301**, 417–432.
26. Krmpotic-Nemanic, J., Kostovic, I., Kelovic, Z. & Nemanic, D. (1980) *Acta Otolaryngol.* **89**, 388–392.
27. Robertson, R.T., Mostamand, F., Kageyama, G.H., Gallardo, K.A. & Yu, J. (1991) *Dev. Brain Res.* **58**, 81–95.
28. Bear, M.F. & Winger, W. (1986) *Nature (London)* **320**, 172–176.
29. Weinberger, N.M. (1995) *Annu. Rev. Neurosci.* **18**, 129–158.
30. Weinberger, N.M. (1998) *Neurobiol. Learn. Mem.* **70**, 226–251.
31. Bjordahl, T.S., Dimyan, M.A. & Weinberger, N.M. (1998) *Behav. Neurosci.* **112**, 467–479.
32. Recanzone, G.H., Guard, D.C. & Phan, M.L. (2000) *J. Neurophysiol.* **83**, 2315–2331.
33. Fitzpatrick, K.A. & Imig, T.J. (1980) *J. Comp. Neurol.* **192**, 589–610.
34. Casagrande, V.A. & Kaas, J.H. (1994) in *Cerebral Cortex*, eds. Peters, A. & Rockland, K. (Plenum, New York), Vol. 10, pp. 201–259.
35. Krubitzer, L.A. & Kaas, J.H. (1990) *J. Neurosci.* **10**, 952–974.
36. Stepniewska, I., Preuss, T.M. & Kaas, J.H. (1993) *J. Comp. Neurol.* **330**, 238–271.
37. Galaburda, A.M. & Sanides, F. (1980) *J. Comp. Neurol.* **190**, 597–610.
38. Rivier, F. & Clarke, S. (1997) *Neuroimage* **6**, 288–304.
39. Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.F. & Pernier, J. (1995) *Electroencephalogr. Clin. Neurophysiol.* **96**, 143–156.
40. Howard, M.A., Volkov, I.O., Abbas, P.J., Damasio, H., Ollendieck, M.C. & Granner, M.A. (1996) *Brain Res.* **724**, 260–264.
41. Elberling, C., Bak, C., Kofoed, B., Lebeck, J. & Saermark, K. (1982) *Scand. Audiol.* **11**, 61–65.
42. Romani, G.L., Williamson, S.J. & Kaufman, L. (1982) *Science* **216**, 1339–1340.
43. Hoke, E.S., Ross, B. & Hoke, M. (1998) *NeuroReport* **9**, 3065–3068.
44. Pantev, C., Hoke, M., Lehnertz, K., Lutkenhoner, B., Anogianakis, G. & Wittkowski, W. (1988) *Electroencephalogr. Clin. Neurophysiol.* **69**, 160–170.
45. Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G. & Elbert, T. (1995) *Electroencephalogr. Clin. Neurophysiol.* **94**, 26–40.
46. Bertrand, O., Perrin, F. & Pernier, J. (1991) *Acta Otolaryngol.* **S491**, 116–123.
47. Yamamoto, T., Uemura, Y. & Llinas, R. (1992) *Acta Otolaryngol.* **112**, 201–204.
48. Tiitinen, H., Alho, K., Huottilainen, R., Ilmoniemi, R.J., Simola, J. & Naatanen, R. (1993) *Psychophysiology* **30**, 537–540.
49. Huottilainen, M., Tiitinen, H., Lavikainen, J., Ilmoniemi, R.J., Pekkonen, E., Sinkkonen, J., Laine, P. & Naatanen, R. (1995) *NeuroReport* **6**, 841–844.
50. Langner, G., Sams, M., Heil, P. & Schultze, H. (1997) *J. Comp. Physiol. A* **181**, 665–676.
51. Lutkenhoner, B. & Steinstrater, O. (1998) *Audiol. Neurootol.* **3**, 191–213.
52. Rosburg, T., Kreitschmann-Andermahr, I., Emmerich, E., Nowak, H. & Sauer, H. (1998) *Neurosci. Lett.* **258**, 105–108.
53. Lauter, J.L., Herscovitch, P., Formby, C. & Raichle, M.E. (1985) *Hear. Res.* **2**, 199–205.
54. de Rossi, G., Paludetti, G., di Nardo, W., Calcagni, M.L., di Guida, D., Almadori, G. & Galli, J. (1996) *Nuklearmedizin* **35**, 112–115.
55. Ottaviani, F., di Girolamo, S., Briglia, G., de Rossi, G., di Guida, D. & di Nardo, W. (1997) *Audiology* **36**, 241–248.
56. Lockwood, A.H., Salvi, R.J., Coad, M.L., Arnold, S.A., Wack, D.S., Murphy, B.W. & Burkard, R.F. (1999) *Cereb. Cortex* **9**, 65–76.
57. Strainer, J.C., Ulmer, J.L., Yetkin, F.Z., Houghton, V.M., Daniels, D.L. & Millen, S.J. (1997) *Am. J. Neuroradiol.* **18**, 601–610.
58. Wessinger, C.M., Buonocore, M.H., Kussmaul, C.L. & Mangun, G.R. (1997) *Hum. Brain Mapp.* **5**, 18–25.
59. Bilecen, D., Scheffler, K., Schmid, N., Tschopp, K. & Seelig, J. (1998) *Hear. Res.* **126**, 19–27.
60. Yang, Y., Engelen, A., Engelen, W., Xu, S., Stern, E. & Silbersweig, D.A. (2000) *Magn. Reson. Med.* **43**, 185–190.
61. Rauschecker, J.P., Tian, B. & Hauser, M. (1995) *Science* **268**, 111–114.
62. Yeterian, E.H. & Pandya, D.N. (1989) *J. Comp. Neurol.* **282**, 80–97.
63. Kosmal, A., Malinowska, M. & Kowalska, D.M. (1997) *Acta Neurobiol. Exp. (Warsz.)* **57**, 165–188.
64. Romanski, L.M., Giguere, M., Bates, J.F. & Goldman-Rakic, P.S. (1997) *J. Comp. Neurol.* **379**, 313–332.

65. Cusick, C.G. (1997) in *Cerebral Cortex, vol. 12*, eds. Rockland, K., Kaas, J.H. & Peters, A. (Plenum, New York), pp. 435–468.
66. Ghosh, S. & Gattera, R. (1995) *Somatosens. Mot. Res.* **12**, 359–378.
67. Wise, S.P., Boussaoud, D., Johnson, P.B. & Caminiti, R. (1997) *Annu. Rev. Neurosci.* **20**, 25–42.
68. Kaas, J.H. & Hackett, T.A. (1998) *Audiol. Neurootol.* **3**, 73–85.
69. Romanski, L.M., Bates, J.F. & Goldman-Rakic, P.S. (1999a) *J. Comp. Neurol.* **403**, 141–157.
70. Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S. & Rauschecker, J.P. (1999b) *Nat. Neurosci.* **2**, 1131–1136.
71. Fuster, J.M. (1995) in *Epilepsy and the Functional Anatomy of the Frontal Lobe*, eds. Jasper, H.H., Riggio, S. & Goldman-Rakic, P.S. (Raven, New York), pp. 9–20.
72. Schall, J.D. (1997) in *Cerebral Cortex*, eds. Rockland, K., Kaas, J.H. & Peters, A. (Plenum, New York), Vol. 12, pp. 527–637.
73. Ito, S.-I. (1982) *Brain Res.* **247**, 39–47.
74. Niki, H., Sugita, S. & Watanabe, M. (1990) in *Vision, Memory, and the Temporal Lobe*, eds. Iwai, E. & Mishkin, M. (Elsevier, Amsterdam), pp. 295–304.
75. Watanabe, M. (1992) *Exp. Brain Res.* **89**, 233–247.
76. Bodner, M., Kroger, J. & Fuster, J.M. (1996) *NeuroReport* **7**, 1905–1908.
77. Mishkin, M. & Manning, F.J. (1978) *Brain Res.* **143**, 313–324.
78. Thorpe, S.J., Rolls, E.T. & Maddison, S. (1983) *Exp. Brain Res.* **49**, 93–115.
79. Gaffan, D. & Murray, E.A. (1990) *J. Neurosci.* **10**, 3479–3493.
80. Schoenbaum, G., Chiba, A.A. & Gallagher, M. (1998) *Nat. Neurosci.* **1**, 155–159.
81. Tremblay, L. & Schultz, W. (1999) *Nature (London)* **398**, 704–708.
82. Benevento, L.A., Fallon, J.H., Davis, B.J. & Rezak, M. (1977) *Exp. Neurol.* **57**, 849–872.
83. Kaas, J.H. & Hackett, T.A. (1999) *Nat. Neurosci.* **2**, 1045–1047.
84. Petersen, S.E., Fox, P.T., Posner, M.I., Mintum, M. & Raichle, M.E. (1988) *Nature (London)* **231**, 585–589.
85. Liegeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V. & Chauvel, P. (1998) *Brain* **121**, 1853–1867.
86. Liegeois-Chauvel, C., Musolino, A., Badier, J., Marquis, P. & Chauvel, P. (1994) *Electroencephalogr. Clin. Neurophysiol.* **92**, 204–214.
87. Price, C. Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K. & Frackowiak, R. (1992) *Neurosci. Lett.* **146**, 179–182.
88. Demonet, J.F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.L., Wise, R., Rascol, A. & Frackowiak, R. (1992) *Brain* **115**, 1753–1768.
89. Zatorre, R.J., Evans, A.C., Meyer, E. & Gjedde, A. (1992) *Science* **256**, 846–849.
90. Binder, J.R., Rao, S.M., Hammeke, T.A., Yetkin, F.Z., Jewmanowicz, A., Bandettini, P.A., Wong, E.C., Estkowski, L.D., Goldstein, M.D., Haughton, V.M. & Hyde, J.S. (1994) *Ann. Neurol.* **35**, 662–672.
91. Berry I., Demonet J.-F., Warach, S., Viillard, G., Boulanouar, K., Franconi, J.M., Marc-Vergnes, J.-P., Edelman, R. & Manelfe, C. (1995) *Neuroimage* **2**, 215–219.
92. Hickcock, G., Love, T., Swinney, D., Wong, E.C. & Buxton, R.B. (1997) *Brain Lang.* **58**, 197–201.
93. Gaschler-Markefski, B., Baumgart, F., Tempelmann, C., Woldorff, M.G. & Scheich, H. (1998) *Neural Plast.* **6**, 69–75.
94. Griffiths, T.D., Rees, G., Rees, A., Green, G.G., Witton, C., Rowe, D., Buchel, C., Turner, R. & Frackowiak, R.S. (1998a) *Nat. Neurosci.* **1**, 74–79.
95. Griffiths, T.D., Buchel, C., Frackowiak, R.S. & Patterson, R.D. (1998b) *Nat. Neurosci.* **1**, 422–427.
96. Scheich, H., Baumgart, F., Gaschler-Markefski, B., Tegeler, C., Tempelmann, C., Heinze, H.J., Schindler, F. & Stiller, D. (1998) *Eur. J. Neurosci.* **10**, 803–809.
97. Baumgart, F., Gaschler-Markefski, B., Woldorff, M.G., Heinze, H.J. & Scheich, H. (1999) *Nature (London)* **400**, 724–726.
98. Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. & Kubo, T. (1999) *Nature (London)* **397**, 116.
99. Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C. & Pike, B. (1999) *Neuroimage* **10**, 417–429.
100. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. & Pike, B. (2000) *Nature (London)* **403**, 309–312.
101. Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J.P., Berry, I., Nespoulous, J.L. & Chollet, F. (1999) *Neuroimage* **9**, 135–144.
102. Jancke, L., Mirzazade, S. & Shah, N.J. (1999) *Neurosci. Lett.* **266**, 125–128.
103. Mummery, C.J., Ashburner, J., Scott, S.K. & Wise, R.J. (1999) *J. Acoust. Soc. Am.* **106**, 449–457.
104. Howard, M.A., Volkov, I.O., Mirsky, R., Garell, P.C., Noh, M.D., Granner, M., Damasio, H., Steinschneider, M., Reale, R.A., Hind, J.E. & Brugge, J.F. (2000) *J. Comp. Neurol.* **416**, 79–92.
105. Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N. & Possing, E.T. (2000) *Cereb. Cortex* **10**, 512–528.

MECHANISMS AND STREAMS FOR PROCESSING OF “WHAT” AND “WHERE” IN AUDITORY CORTEX

Josef P. Rauschecker* and Biao Tian

Georgetown Institute for Cognitive and Computational Sciences, Washington, DC 20007

The functional specialization and hierarchical organization of multiple areas in rhesus monkey auditory cortex were examined with various types of complex sounds. Neurons in the lateral belt areas of the superior temporal gyrus were tuned to the best center frequency and bandwidth of band-passed noise bursts. They were also selective for the rate and direction of linear frequency modulated sweeps. Many neurons showed a preference for a limited number of species-specific vocalizations (“monkey calls”). These response selectivities can be explained by nonlinear spectral and temporal integration mechanisms. In a separate series of experiments, monkey calls were presented at different spatial locations, and the tuning of lateral belt neurons to monkey calls and spatial location was determined. Of the three belt areas the anterolateral area shows the highest degree of specificity for monkey calls, whereas neurons in the caudolateral area display the greatest spatial selectivity. We conclude that the cortical auditory system of primates is divided into at least two processing streams, a spatial stream that originates in the caudal part of the superior temporal gyrus and projects to the parietal cortex, and a pattern or object stream originating in the more anterior portions of the lateral belt. A similar division of labor can be seen in human auditory cortex by using functional neuroimaging.

The visual cortex of nonhuman primates is organized into multiple, functionally specialized areas (1, 2). Among them, two major pathways or “streams” can be recognized that are involved in the processing of object and spatial information (3). Originally postulated on the basis of behavioral lesion studies (4), these “what” and “where” pathways both originate in primary visual cortex V1 and are, respectively, ventrally and dorsally directed. Already in V1 neurons are organized in a domain-specific fashion, and separate pathways originate from these domains before feeding into the two major processing streams (5). Neurons in area V4, which is part of the “what” pathway or ventral stream, are highly selective for the color and size of visual objects (6, 7) and, in turn, project to inferotemporal areas containing complex visual object representations (8, 9). Neurons in area V5 (or MT), as part of the “where” pathway or dorsal stream, are highly selective for the direction of motion (10) and project to the parietal cortex, which is crucially involved in visual spatial processing (11–13). Both pathways eventually project to prefrontal cortex, where they end in separate target regions (14) but may finally converge (15). A similar organization has been reported recently for human visual cortex on the basis of neuroimaging studies (16, 17).

Compared with this elaborate scheme that has been worked out for visual cortical organization, virtually nothing has been known about the functional organization of higher auditory cortical pathways, even though a considerable amount of anatomical information had been collected early on (18–23). Around the same time, initial electrophysiological single-unit mapping studies with tonal stimuli were also undertaken (24). These studies described several areas on the supratemporal plane within rhesus monkey auditory cortex. Primary auditory cortex A1 was found to be surrounded by several other auditory areas. A rostralateral area (RL, later renamed rostral area, R) shares its low-frequency border with A1, whereas a caudomedial area (CM) borders A1 at its high-frequency end. All three of these areas are tonotopically organized and mirror-symmetric to each other along the frequency axis. In addition, medial and lateral regions were reported as responsive to auditory stimuli but could not be characterized further with tonal stimuli.

Organization of Thalamocortical Auditory Pathways

Interest in the macaque’s auditory cortical pathways was revived with the advent of modern histochemical techniques in combination with the use of tracers to track anatomical connections (25–28). Injection of these tracers into physiologically identified and characterized locations further strengthens this approach. Thus, after determining the tonotopic maps on the supratemporal plane with single-unit techniques, three different tracers were injected into identical frequency representations in areas A1, R, and CM (29). As a result of these injections, neurons in the medial geniculate nucleus (MGN) became retrogradely labeled. Label from injections into A1 and R was found in the ventral division of the MGN, which is the main auditory relay nucleus, whereas injections into CM labeled only the dorsal and medial divisions. This means that A1 and R both receive input from the ventral part of the MGN in parallel, whereas CM does not.

As a consequence, making lesions in primary auditory cortex A1 has different effects on responses in areas R and CM (29). When auditory responses in area R of the same animal before and after the A1 lesion were compared, they were essentially unchanged. By contrast, auditory responses in area CM, especially those to pure tones, were virtually abolished after the lesion. Thus, area R seems to receive its input independently of A1, whereas CM responses do depend on the integrity of A1. In other words, the rhesus monkey auditory system, beginning at the level of the medial geniculate nucleus (or even the cochlear nuclei), is organized both serially and in parallel, with A1 and R both receiving direct input from the ventral part of the medial geniculate nucleus.

As part of the mapping studies of the supratemporal plane, numerous examples of spatially tuned neurons were discovered in area CM (30), although this was not systematically explored

*To whom reprint requests should be addressed at: Georgetown Institute for Cognitive and Computational Sciences, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007. E-mail: rauscheckerj@giccs.georgetown.edu.

This paper was presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: R, rostral area; CM, caudomedial area; AL, anterolateral area; ML, middle lateral area; CL, caudolateral area; MGN, medial geniculate nucleus; STG, superior temporal gyrus; BPN, band-passed noise; FM, frequency modulated; MC, monkey call; PET, positron-emission tomography.

at that time. However, Leinonen *et al.* (31) had earlier described auditory spatial tuning in neurons of area Tpt of Pandya and Sanides (19), which is adjacent to or overlapping with CM. The lateral areas receiving input from both A1 and R, on the other hand, may be the beginning of an auditory pattern or object stream: as we will see in the following, these areas contain neurons responsive to species-specific vocalizations and other complex sounds.

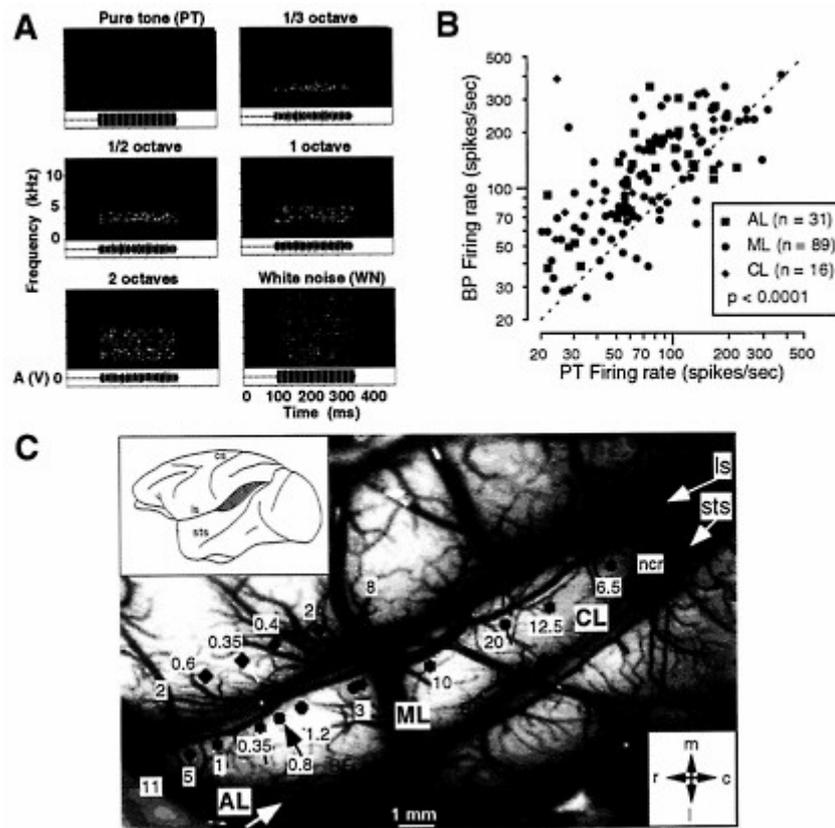


Fig. 1. Mapping of lateral belt areas in the rhesus monkey. (A) Band-passed noise (BPN) bursts of various bandwidths and constant center frequency are displayed as spectrograms. (B) Scattergram comparing responses to BPN and pure-tone (PT) stimuli in the same neurons. (C) Reconstruction of best center frequency maps showing cochleotopic organization of anterolateral (AL), middle lateral (ML), and caudolateral (CL) areas on the STG of one monkey (32). BFc, best center frequency; cs, central sulcus; ls, lateral sulcus; sts, superior temporal sulcus; ncr, no clear response.

Use of Complex Sound Stimuli in Neurons of the Lateral Belt Band-Passed Noise (BPN) Bursts. From the earlier comparison with the visual system it becomes clear almost immediately that stimulation of neurons in higher visual areas with small spots of light is doomed to failure, because neurons integrate over larger areas of the sensory epithelium. By analogy, neurons in higher auditory areas should not and will not respond well to pure tones of a single frequency. The simplest step toward designing effective auditory stimuli for use on higher-order neurons is, therefore, to increase the bandwidth of the sound stimuli. BPN bursts centered at a specific frequency (Fig. 1A) are the equivalent of spots or bars of light in the visual system. It turns out that such stimuli are indeed highly effective in evoking neuronal responses from lateral belt (Fig. 1B). Virtually every neuron can now be characterized, and a tonotopic (or better cochleotopic) organization becomes immediately apparent (Fig. 1C). On the basis of reversals of best center frequency within these maps, three lateral belt areas can be defined, which we termed anterolateral (AL), middle lateral (ML), and caudolateral (CL) areas (32). These three areas are situated adjacent to and in parallel with areas R, A1, and CM, respectively. The availability of these lateral belt areas for detailed exploration brings with it the added bonus that they are situated on the exposed surface of the superior temporal gyrus (STG), an advantage that should not be underestimated. A parabelt region even more lateral and ventral on the STG has been defined on histochemical grounds (25, 33, 34), but belt and parabelt regions have not yet been distinguished electrophysiologically.

Responses of lateral belt neurons to BPN bursts are usually many times stronger than responses to pure tones (Fig. 1B). Facilitation of several hundred percent is the rule. Furthermore, the neurons often respond best to a particular bandwidth of the noise bursts, a property referred to as bandwidth tuning (32), and the response does not increase simply with bandwidth. Only few neurons respond less well to the BPN bursts than to the pure tones. The peak in the bandwidth tuning curve was generally unaffected by changes in intensity of the stimulus, which is important if one considers the involvement of such neurons in the decoding of auditory patterns. Best bandwidth seems to vary along the mediolateral axis of the belt areas, orthogonally to the best center frequency axis.

Frequency-Modulated (FM) Sweeps. Adding temporal complexity to a pure-tone stimulus creates an FM sweep. Neurons in the lateral belt respond vigorously to linear FM sweeps and are highly selective to both their rate and direction (30, 35, 36). Selectivity to FM is already found in primary auditory cortex or even the inferior colliculus (37–40), but is even more pronounced in the lateral belt. FM selectivity differs significantly between areas on the lateral belt (35), with AL neurons responding better to slow FM rates (≈ 10 kHz/s) and neurons in CL responding best to high rates (≈ 100 kHz/s).

It is attractive to use FM sweeps as stimuli in the auditory cortex for another reason: as argued previously (35), FM sweeps are equivalent to moving light stimuli, which have proven so highly effective for neurons of the visual cortex (41). The comparable selectivity in both sensory modalities suggests that cortical modules across different areas could apply the same temporal-spatial algorithm onto different kinds of input.

Monkey Vocalizations. A third class of complex sounds that we have used extensively for auditory stimulation in the lateral belt are vocalizations from the rhesus monkey's own repertoire. Digitized versions of such calls, recorded from free-ranging monkeys on the island of Cayo Santiago, were available to us from a sound library assembled by Marc Hauser at Harvard University. Hauser (42) classifies rhesus monkey vocalizations into roughly two dozen different categories, which can be subdivided phonetically into three major groups: tonal, harmonic, and noisy calls (35). Tonal calls are characterized by their concentration of energy into a narrow band of frequencies that can be modulated over time. Harmonic calls contain large numbers of higher harmonics in addition to the fundamental frequency. Noisy calls, often uttered in aggressive social situations, are characterized by broadband frequency spectra that are temporally modulated. The semantics of these calls have been studied extensively (42).

A standard battery of seven different calls, which were representative on the basis of both phonetic and semantic properties, was routinely used for stimulation in single neurons of the lateral belt. A great degree of selectivity of neuronal responses was found for different types of calls. Despite a similar bandwidth of some of the calls, many neurons respond better to one type of call than to another, obviously because of the different fine structure in both the spectral and the temporal domain of the different calls. A preference index (PI) was established by measuring the peak firing rate in response to each of the seven calls and counting the number of calls that elicit a significant increase in firing rate. PI 1 refers to neurons that respond only to a single call, an index of 7 refers to neurons that respond to all seven calls, and indices of 2 to 6 refer to the corresponding numbers in between. The PI distribution in most animals reveals that few neurons respond only to a single call, few respond to all seven calls, but most respond to a number in between, usually 3, 4, or 5 of the calls. This suggests that the lateral belt areas are not yet the end stage of the pathway processing monkey vocalizations. Alternatively, monkey calls (MCs) could be processed by a network of neurons rather than single cells, a suggestion that is of course not mutually exclusive with the first.

Spectral and Temporal Integration in Lateral Belt Neurons

The next step in our analysis was to look for the mechanisms that make neurons in the lateral belt selective for certain calls rather than others. One pervasive mechanism that we found was "spectral integration." MCs can, for example, be broken down into two spectral components, a low-pass and a high-pass filtered version. The neuron in the example of Fig. 2A, which responded well to the total call with the full spectrum, did not respond as well to the low-pass-filtered version and not at all to the high-pass-filtered version. When both spectral components were combined again, the response was restored to the full extent. Thus, neurons in the lateral belt combine information from different frequency bands, and the nonlinear combination of information in the spectral domain leads to response selectivity. In some instances, however, suppression instead of facilitation of the response by combining two spectral components was also found.

A similar combination sensitivity is found in the time domain. If an MC with two syllables is used for stimulation, it is often found that the response to each syllable alone is minimal, whereas the combination of the two syllables in the right temporal order leads to a large response (Fig. 2B). Temporal integration will have to occur within a window as long as several hundreds of milliseconds. The neural and synaptic mechanisms that can implement such integration times have yet to be clarified.

Neurons of the kind reported here are much more frequently found in lateral belt than in A1 ($P < 0.001$). They also must be at a higher hierarchical level than the types of bandpass-selective neurons described earlier. Combining inputs from different frequency bands ("formants") or with different time delays in a specific way could thus lead to the creation of certain types of "call detectors" (35). Similar models have been suggested for song selectivity in songbirds (43) and selectivity to specific echo delay combinations in bats (44).

Origins of "What" and "Where" Streams in the Lateral Belt

In our next study we compared the response selectivity of single neurons in the lateral belt region of rhesus macaque auditory cortex simultaneously to species-specific vocalizations and spatially localized sounds (45). The purpose of this study was not only to learn more about the functional specialization of the lateral belt areas, but more specifically to test the hypothesis that the cortical auditory system divides into two separate streams for the processing of "what" and "where."

After mapping the lateral belt areas in the usual manner, using BPN bursts centered at different frequencies (32), the same set of MCs as in our previous studies was used to determine the selectivity of the neurons for MCs. A horizontal speaker array was used to test the spatial tuning of the same neurons in 20°-steps of azimuth. To determine a neuron's selectivity for both MCs and spatial position, 490 responses were evaluated quantitatively in every neuron.

The results of the study from a total of 170 neurons in areas AL, ML, and CL can be summarized as follows (Fig. 3):

- (i) Greatest spatial selectivity was found in area CL.
- (ii) Greatest selectivity for MCs was found in area AL.
- (iii) In CL, monkey call selectivity often covaried with spatial selectivity.

In terms of processing hierarchies in rhesus monkey auditory cortex the following can be determined:

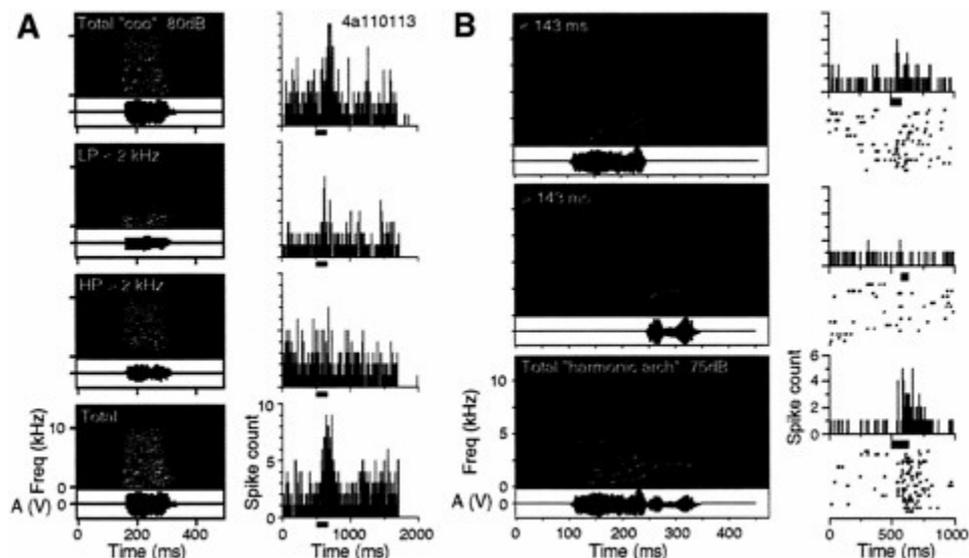
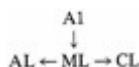


Fig. 2. Spectral and temporal integration in single neurons of the lateral belt in primates. Digitized monkey vocalizations were presented as stimuli, either as complete calls or as components manipulated in the spectral or temporal domain. (A) Nonlinear spectral summation. The “coo” call consists of a number of harmonic components and elicits a good response. If the call is low-pass-filtered with a cutoff frequency of 2 kHz, a much smaller response is obtained. The same is true for the high-pass-filtered version. Stimulation with the whole signal is repeated to demonstrate the reliability of the result (*Bottom*) (32). (B) Nonlinear temporal summation. The “harmonic arch” call consists of two “syllables.” Each of them alone elicits a much smaller response than the complete call.

Spatial selectivity increases from ML→CL, and is lowest in AL.



MC selectivity increases from ML→AL, but also from ML→CL.



We can conclude, therefore, that the caudal belt region is the major recipient of auditory spatial information from A1 (and subcortical centers). This spatial information is relayed from the caudal belt to posterior parietal cortex and to dorsolateral prefrontal cortex (46). The caudal part of the STG (areas CL and CM) can thus be considered the origin of a “where” -stream for auditory processing. The anterior areas of the STG, on the other hand, are major recipients of information relevant for auditory object or pattern perception. Projecting on to orbitofrontal cortex, they can thus be thought of as forming an auditory “what” -stream (46). As can be seen in the following section, recent results of human imaging strongly support this view. However, more traditional theories of speech perception have emphasized a role of posterior STG in phonological decoding. It is important to note, therefore, that selectivity for communication signals is also relayed to the caudal STG, where it is combined with information about the localization of sounds (see *Comparison of Monkey and Human Data* below for further discussion).

“What” and “Where” in Human Auditory Cortex

Processing of Speech-Like Stimuli in the Superior Temporal Cortex. How can research on nonhuman primates be relevant to the understanding of human speech perception? First, there is a striking resemblance of the spectrotemporal phonetic structure of human speech sounds to those of other species-specific vocalizations (35). Looking at human speech samples, one can recognize BPN portions contained in the coding of different types of consonants, e.g., fricatives or plosives. In addition, the presence of FM sweeps in certain phonemes and formant transitions is noticeable. Fast FM sweeps are critical for the encoding and distinction of consonant/vowel combinations such as “ba,” “da,” and “ga” (47, 48). It appears more than likely that human speech sounds are decoded by types of neurons similar to the ones found in macaque auditory cortex, perhaps with even finer tuning to the relevant parameter domains.

Second, there are intriguing similarities between the two species in terms of anatomical location. The STG in humans has been known for some time to be involved in the processing of speech or phonological decoding. This evidence stems from a number of independent lines of investigation. Lesions of the STG by stroke lead to sensory aphasia (49) and word deafness (50). Electrical stimulation of the cortical surface in the STG leads to temporary “functional lesions” used during presurgical screening in epileptic patients (51). Using this approach, it can be shown that the posterior superior temporal region is critically involved in phoneme discrimination. In addition, Zatorre *et al.* (52) have shown, using positron-emission tomography (PET) techniques, that the posterior superior temporal region lights up with phoneme stimulation.

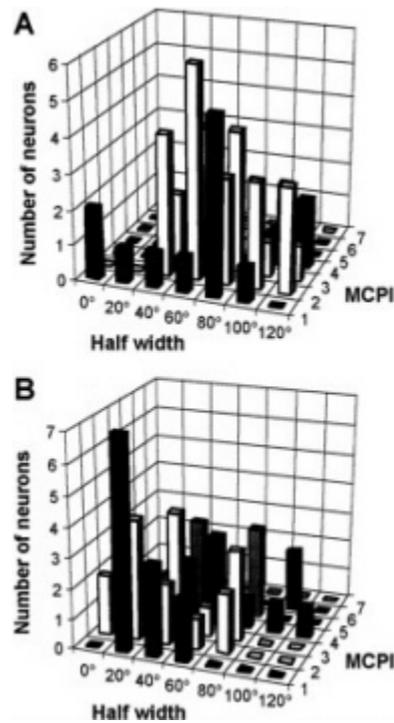


Fig. 3. Monkey-call preference index (MCPI) and spatial half width in the same neurons of rhesus monkey lateral belt. Results from the AL and CL areas are plotted in *A* and *B*, respectively.

Using techniques of functional magnetic resonance imaging (MRI), we are able to map the activation of auditory cortical areas directly in the human brain. Functional MRI gives much better resolution than PET and is therefore capable of demonstrating such functional organizational features as tonotopic organization (53). The same types of comparisons as in our monkey studies were used and clearly demonstrate that pure tones activate only limited islands within the core region of auditory cortex on the supratemporal plane, whereas BPN stimulation leads to more extensive zones of activation, particularly in the lateral belt. Several discrete cortical maps can be discerned with certainty, and they correspond to the maps identified in monkey auditory cortex (35, 54).

In a next step, the comparison of BPN stimulation with stimulation by consonant/vowel (CV) combinations shows that the latter leads to yet more extensive activation. Different CV tokens (corrected for fundamental frequency) lead to distinct but overlapping activations (Fig. 4), which suggest the existence of a “phonetic mapping.” In some subjects, stimulation with phonemes leads to asymmetry between the left and right hemispheres, with CV combinations often leading to more prominent activation on the left, although this is not always consistently the case (55, 56).

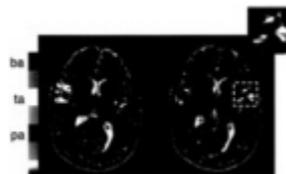


Fig. 4. Functional MRI study of the STG in a human subject while listening to speech sounds. A phonemic map may be recognized anterior of Heschl's gyrus resulting from superposition of activation by three different consonant/ vowel combinations [ba, da, ga (64)[†]; courtesy of Brandon Zielinski, Georgetown University].

One should also emphasize that experimental set-up and instructions are critical for the success of functional MRI studies. Asking subjects to pay specific attention to target words within a list of words presented to them during the functional MRI scan greatly enhances the signal that can be collected as compared with simple passive listening (57).

Auditory Spatial Processing in the Inferior Parietal Cortex. Evidence for the existence of a dorsal stream in humans, as in the visual system, having to do with the processing of spatial information comes from PET studies of activation by virtual auditory space stimuli (58, 59). In these studies, the stimuli were generated by a computer system (Tucker-Davies Technology, Gainesville, FL, Power Dac PD1) based on head-related transfer functions established by Wightman and Kistler (60), and presented by headphones. The sounds had characteristic interaural time and intensity differences, as well as spectral cues encoding different azimuth locations. Use of these stimuli led to specific activation in a region of the inferior parietal lobe (Fig. 5A). The latter is normally associated with spatial analysis of visual patterns (61). However, the activation by auditory spatial analysis led to a focus that was about 8 mm more inferior than the location usually found by visual stimuli. Furthermore, when visual and auditory space stimuli were tested in the same subjects, there were clearly distinct activation foci stemming from the two modalities (59).

There was also a slight bias toward activation in the right hemisphere, which is consistent with the idea that the right hemisphere is more involved in spatial analysis than the left hemisphere. At the same time as we received activation of inferior parietal areas by virtue of auditory space stimuli we also got a *de*-activation in temporal areas bilaterally (Fig. 5B), which supports the idea that these areas are involved in auditory tasks other than spatial ones—for example, those involved in the decoding of auditory patterns, including speech.

Comparison of Monkey and Human Data

In comparing monkey and human data, one apparent paradox may be noted: Speech perception in humans is traditionally associated with the posterior portion of the STG region, often referred to as “Wernicke's area.” In rhesus monkeys, on the other hand, we and others (31, 62) have found neurons in this region (areas Tpt, CM, and CL) that are highly selective for the

[†]Zielinski, B.A., Liu, G., and Rauschecker, J.P. (2000) *Soc. Neurosci. Abstr.* **26**, 737.3 (abstr.).

spatial location of sounds in free field, which suggests a role in auditory localization. Neurons in the anterior belt regions, on the other hand, are most selective for MCs. Several explanations for this paradox, which are not mutually exclusive, appear possible: (i) Speech processing in humans may be localized not only in posterior but also in anterior STG. Evidence for this comes from recent imaging studies (56, 63). (ii) In evolution, the anterior part of the temporal lobe may have grown disproportionately, as has also been argued with regard to prefrontal cortex (17). The precursor of Wernicke's area in the monkey may thus be situated relatively more anterior and encompass a good portion of the anterolateral belt (area AL) or even more anterior regions of the monkey's STG. (iii) Spatial selectivity in the caudal belt may play a dual role in sound localization as well as identification of sound sources on the basis of location ("mixture party effect"). Hypothetically, its medial portion (CM) may be more specialized in localization (62) than its lateral portion (CL).

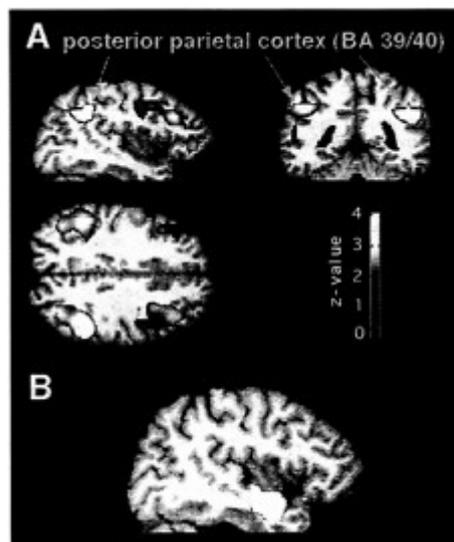


Fig. 5. PET activation of the human brain during localization of virtual auditory space stimuli. (A) Statistical parametric mapping (SPM) projections of significant areas of activation from sagittal, coronal, and axial directions. PET areas are superimposed onto representative MRI sections. (B) Area of deactivation in the right anterior STG caused by auditory localization. [Based on data from Weeks *et al.* (58).]

Conclusions

In summary, we have collected evidence from studies in nonhuman as well as human primates that the auditory cortical pathways are organized in parallel as well as serially. The lateral belt areas of the STG seem to be critically involved in the early processing of species-specific vocalizations as well as human speech. By contrast, a pathway originating from the caudal or caudomedial part of the supratemporal plane and involving the inferior parietal areas seems to be an important way station for the processing of auditory spatial information (Fig. 6). As we have emphasized before, it is important that investigations in human and nonhuman primates continue concurrently, using both functional brain imaging techniques noninvasively in humans and microelectrode studies in nonhuman primates. Although direct homologies between the two species have to be

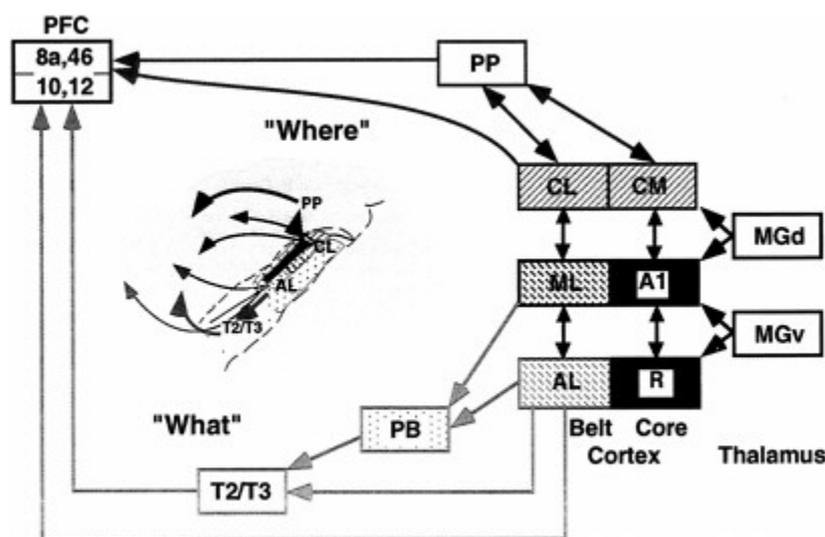


Fig. 6. Schematic flow diagram of "what" and "where" streams in the auditory cortical system of primates. The ventral "what"-stream is shown in green, the dorsal "where"-stream, in red. [Modified and extended from Rauschecker (35); prefrontal connections (PFC) based on Romanski *et al.* (46).] PP, posterior parietal cortex; PB, parabelt cortex; MGd and MGv, dorsal and ventral parts of the MGN.

drawn with care, only the combination of both techniques can eventually reveal the mechanisms and functional organization of higher auditory processing in humans and lead to effective therapies for higher speech disorders.

We deeply appreciate the collaborative spirit of Patricia Goldman-Rakic (Yale University), Mark Hallett (National Institute of Neurological Disorders and Stroke), Marc Hauser (Harvard University), and Mortimer Mishkin (National Institute of Mental Health). The help of the following individuals is acknowledged gratefully: Jonathan Fritz (National Institute of Mental Health), Khalaf Bushara and Robert Weeks (National Institute of Neurological Disorders and Stroke), Liz Romanski (Yale University), and Amy Durham, Alexander Kustov, Aaron Lord, David Reser, Jenny VanLare, Mark Wessinger, and Brandon Zielinski (all Georgetown University). This work was supported by Grants R01 DC 03489 from the National Institute on Deafness and Other Communicative Disorders and DAMD17-93-V-3018 from the U.S. Department of Defense.

1. Zeki, S.M. (1978) *Nature (London)* **274**, 423–428.
2. DeYoe, E.A. & Van Essen, D.C. (1988) *Trends Neurosci.* **11**, 219–226.
3. Ungerleider, L.G. & Mishkin, M. (1982) in *Analysis of Visual Behaviour*, eds. Ingle, D.J., Goodale, M.A. & Mansfield, R.J.W. (MIT Press, Cambridge, MA), pp. 549–586.
4. Mishkin, M., Ungerleider, L.G. & Macko, K.A. (1983) *Trends Neurosci.* **6**, 414–417.
5. Livingstone, M.S. & Hubel, D.H. (1988) *Science* **240**, 740–749.
6. Zeki, S. (1983) *Neuroscience* **9**, 741–765.
7. Desimone, R. & Schein, S.J. (1987) *J. Neurophysiol.* **57**, 835–868.
8. Desimone, R. (1991) *J. Cogn. Neurosci.* **3**, 1–8.
9. Tanaka, K. (1997) *Curr. Opin. Neurobiol.* **7**, 523–529.
10. Movshon, J.A. & Newsome, W.T. (1996) *J. Neurosci.* **16**, 7733–7741.
11. Colby, C.L. & Goldberg, M.E. (1999) *Annu. Rev. Neurosci.* **22**, 319–349.
12. Andersen, R.A., Snyder, L.H., Bradley, D.C. & Xing, J. (1997) *Annu. Rev. Neurosci.* **20**, 303–330.
13. Sakata, H. & Kusunoki, M. (1992) *Curr. Opin. Neurobiol.* **2**, 170–174.
14. O'Scalaidhe, S.P., Wilson, F.A. & Goldman-Rakic, P.S. (1997) *Science* **278**, 1135–1138.
15. Rao, S.C., Rainer, G. & Miller, E.K. (1997) *Science* **276**, 821–824.
16. Ungerleider, L.G. & Haxby, J.V. (1994) *Curr. Opin. Neurobiol.* **4**, 157–165.
17. Courtney, S.M., Petit, L., Maisog, J.M., Ungerleider, L.G. & Haxby, J.V. (1998) *Science* **279**, 1347–1351.
18. Woolsey, C.N. & Walzl, E.M. (1942) *Bull. Johns Hopkins Hosp.* **71**, 315–344.
19. Pandya, D.N. & Sanides, F. (1973) *Z. Anat. Entwicklungsgesch.* **139**, 127–161.
20. Mesulam, M.M. & Pandya, D.N. (1973) *Brain Res.* **60**, 315–333.
21. Galaburda, A.M. & Pandya, D.N. (1983) *J. Comp. Neurol.* **221**, 169–184.
22. Burton, H. & Jones, E.G. (1976) *J. Comp. Neurol.* **168**, 249–302.
23. Pandya, D.N. & Seltzer, B. (1982) *Trends Neurosci.* **5**, 386–390.
24. Merzenich, M.M. & Brugge, J.F. (1973) *Brain Res.* **50**, 275–296.
25. Morel, A., Garraghty, P.E. & Kaas, J.H. (1993) *J. Comp. Neurol.* **335**, 437–459.
26. Jones, E.G., Dell'Anna, M.E., Molinari, M., Rausell, E. & Hashikawa, T. (1995) *J. Comp. Neurol.* **362**, 153–170.
27. Molinari, M., Dell'Anna, M.E., Rausell, E., Leggio, M.G., Hashikawa, T. & Jones, E.G. (1995) *J. Comp. Neurol.* **362**, 171–194.
28. Kosaki, H., Hashikawa, T., He, J. & Jones, E.G. (1997) *J. Comp. Neurol.* **386**, 304–316.
29. Rauschecker, J.P., Tian, B., Pons, T. & Mishkin, M. (1997) *J. Comp. Neurol.* **382**, 89–103.
30. Rauschecker, J.P. (1997) *Acta Otolaryngol. Suppl.* **532**, 34–38.
31. Leinonen, L., Hyvärinen, J. & Sovijärvi, A.R.A. (1980) *Exp. Brain Res.* **39**, 203–215.
32. Rauschecker, J.P., Tian, B. & Hauser, M. (1995) *Science* **268**, 111–114.
33. Hackett, T.A., Stepniewska, I. & Kaas, J.H. (1998) *J. Comp. Neurol.* **394**, 475–495.
34. Kaas, J.H., Hackett, T.A. & Tramo, M.J. (1999) *Curr. Opin. Neurobiol.* **9**, 164–170.
35. Rauschecker, J.P. (1998) *Audiol Neurootol* **3**, 86–103.
36. Rauschecker, J.P. (1998) *Curr. Opin. Neurobiol.* **8**, 516–521.
37. Whitfield, I.C. & Evans, E.F. (1965) *J. Neurophysiol.* **28**, 655–672.
38. Suga, N. (1968) *J. Physiol. (London)* **198**, 51–80.
39. Suga, N. (1969) *J. Physiol. (London)* **200**, 555–574.
40. Mendelson, J.R. & Cynader, M.S. (1985) *Brain Res.* **327**, 331–335.
41. Hubel, D.H. & Wiesel, T.N. (1962) *J. Physiol. (London)* **160**, 106–154.
42. Hauser, M.D. (1996) *The Evolution of Communication* (MIT Press, Cambridge, MA).
43. Margoliash, D. & Fortune, E.S. (1992) *J. Neurosci.* **12**, 4309–4326.
44. Suga, N., O'Neill, W.E. & Manabe, T. (1978) *Science* **200**, 778–781.
45. Rauschecker, J.P., Durham, A., Kustov, A., Lord, A. & Tian, B. (1999) *Soc. Neurosci. Abstr.* **25**, 394.
46. Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S. & Rauschecker, J.P. (1999) *Nat. Neurosci.* **2**, 1131–1136.
47. Liberman, A.M., Cooper, F.S., Shankweiler, D.P. & Studdert-Kennedy, M. (1967) *Psychol. Rev.* **74**, 431–461.
48. Tallal, P. & Piercy, M. (1973) *Nature (London)* **241**, 468–469.
49. Wernicke, C. (1874) *Der aphasische Symptomenkomplex* (Cohn Weigert, Breslau, Poland).
50. Barrett, A.M. (1910) *J. Nerv. Ment. Dis.* **37**, 73–92.
51. Boatman, D., Lesser, R.P. & Gordon, B. (1995) *Brain Lang.* **51**, 269–290.
52. Zatorre, R.J., Evans, A.C., Meyer, E. & Gjedde, A. (1992) *Science* **256**, 846–849.
53. Wessinger, C.M., Buonocore, M., Kussmaul, C.L. & Mangun, G.R. (1997) *Hum. Brain Mapp.* **5**, 18–25.
54. Wessinger, C.M., Tian, B., VanMeter, J.W., Platenberg, R.C., Pekar, J. & Rauschecker, J.P. (2000) *J. Cognitive Neurosci.*, (in press).
55. Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M. & Prieto, T. (1997) *J. Neurosci.* **17**, 353–362.
56. Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N. & Possing, E.T. (2000) *Cereb. Cortex* **10**, 512–528.
57. Grady, C.L., VanMeter, J.W., Maisog, J.M., Pietrini, P., Krasuski, J. & Rauschecker, J.P. (1997) *Neuroreport* **8**, 2511–2516.
58. Weeks, R.A., Aziz-Sultan, A., Bushara, K.O., Tian, B., Wessinger, C.M., Dang, N., Rauschecker, J.P. & Hallett, M. (1999) *Neurosci. Lett.* **262**, 155–158.
59. Bushara, K.O., Weeks, R.A., Ishii, K., Catalan, M.-J., Tian, B., Rauschecker, J.P. & Hallett, M. (1999) *Nat. Neurosci.* **2**, 759–766.
60. Wightman, F.L. & Kistler, D.J. (1989) *J. Acoust. Soc. Am.* **85**, 858–867.
61. Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P. & Grady, C.L. (1994) *J. Neurosci.* **14**, 6336–6353.
62. Recanzone, G.H., Guard, D.C., Phan, M.L. & Su, T.K. (2000) *J. Neurophysiol.* **83**, 2723–2739.
63. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. & Pike, B. (2000) *Nature (London)* **403**, 309–312.
64. Morad, A., Perez, C.V., Van Lare, J.E., Wessinger, C.M., Zielinski, B.A. & Rauschecker, J.P. (1999) *Neuroimage* **9**, S996.

THE CORTICOFUGAL SYSTEM FOR HEARING: RECENT PROGRESS

Nobuo Suga*, Enquan Gao†, Yunfeng Zhang‡, Xiaofeng Ma, and John F. Olsen

Department of Biology, Washington University, One Brookings Drive, St. Louis, MO 63130

Peripheral auditory neurons are tuned to single frequencies of sound. In the central auditory system, excitatory (or facilitatory) and inhibitory neural interactions take place at multiple levels and produce neurons with sharp level-tolerant frequency-tuning curves, neurons tuned to parameters other than frequency, cochleotopic (frequency) maps, which are different from the peripheral cochleotopic map, and computational maps. The mechanisms to create the response properties of these neurons have been considered to be solely caused by divergent and convergent projections of neurons in the ascending auditory system. The recent research on the corticofugal (descending) auditory system, however, indicates that the corticofugal system adjusts and improves auditory signal processing by modulating neural responses and maps. The corticofugal function consists of at least the following subfunctions. (i) Egocentric selection for short-term modulation of auditory signal processing according to auditory experience. Egocentric selection, based on focused positive feedback associated with widespread lateral inhibition, is mediated by the cortical neural net working together with the corticofugal system. (ii) Reorganization for long-term modulation of the processing of behaviorally relevant auditory signals. Reorganization is based on egocentric selection working together with nonauditory systems. (iii) Gain control based on overall excitatory, facilitatory, or inhibitory corticofugal modulation. Egocentric selection can be viewed as selective gain control. (iv) Shaping (or even creation) of response properties of neurons. Filter properties of neurons in the frequency, amplitude, time, and spatial domains can be sharpened by the corticofugal system. Sharpening of tuning is one of the functions of egocentric selection.

auditory system | descending system | learning and memory plasticity | tonotopic map

The central auditory system creates many physiologically distinct types of neurons for auditory signal processing. Their response properties have been interpreted to be produced by divergent and convergent interactions between neurons in the ascending auditory system. Until recently, the contribution of the descending (corticofugal) system to the shaping (or even creation) of their response properties has hardly been considered. Recent findings indicate that the corticofugal system plays important roles in shaping or even creating the response properties of central auditory neurons and in reorganizing cochleotopic (frequency) and computational (e.g., echo-delay) maps. Therefore, the understanding of the neural mechanisms for auditory signal processing is incomplete without the exploration of the functional roles of the corticofugal system. In this article, we first enumerate several types of neurons and computational maps created in the bat's central auditory system and then describe the anatomy and physiology of the corticofugal system.

Neurons Tuned to Acoustic Parameters Characterizing Biosonar Signals and Cochleotopic and Computational Maps

All peripheral neurons are tuned to single frequencies (best frequencies, BFs). In the central auditory system, excitatory, inhibitory, and facilitatory neural interactions take place at multiple levels and produce neurons with sharp level-tolerant (the width of a frequency tuning curve is narrow regardless of sound levels) frequency tuning curves (1) and also neurons tuned to specific values of parameters other than frequency. Some of these neurons apparently are related to the processing of biosonar signals. They are latency-constant, phasic on-responding neurons (2, 3); paradoxical latency-shift neurons (4); duration-tuned neurons (5); frequency modulation (FM)-sensitive or -specialized neurons (6–8); sharply frequency-tuned neurons showing level tolerance (9, 10); FM or amplitude modulation (AM) rate-tuned neurons (11); constant frequency (CF)/CF and FM-FM combination-sensitive neurons (12); neurons tuned to particular combinations of frequency and amplitude (e.g., ref. 9); and binaural neurons (see ref. 13 for review). The biosonar signals of the mustached bat consist of CF and FM components. CF/CF and FM-FM neurons, respectively, are tuned to specific combinations of CF or FM components of the emitted pulse and its echo.

In the auditory cortex (AC) of the little brown bat, neurons with short best durations are located ventrally to those with long best durations. They may form a duration axis (14). In the AC of the mustached bat, CF/CF neurons are clustered in the CF/CF area and form frequency-vs.-frequency coordinates for the systematic representation of Doppler shift; FM-FM neurons are clustered in the FM-FM area and form an echo-delay axis for the systematic representation of target distance (see ref. 15 for review); and Doppler-shifted constant-frequency (DSCF) neurons are clustered in the DSCF area and form frequency-vs.-amplitude coordinates for the fine spatio-temporal representation of periodic frequency and amplitude modulations of echoes from flying insects (9). In the superior colliculus of the big brown bat, there is a space map, and some neurons are tuned to a sound source at a particular azimuth, elevation, and depth (16). No auditory space map has been found in the AC. Instead, it has been found that two types of binaural neurons (I-E and E-E) form binaural bands in the AC (mustached bats, ref. 17; cats, ref. 18), and that the best azimuth to excite neurons varies systematically along the frequency axis of the AC (mustached bats, ref. 19; big brown bats, ref. 20).

Neural mechanisms to create the response properties of neurons and the computational maps listed above have been explained by various mechanisms, such as inhibition, coincidence

*To whom reprint requests should be addressed. E-mail: suga@biology.wustl.edu.

†Present address: Department of Neurology, General Hospital of the Navy, Beijing, 100037, P.R.China.

‡Present address: Department of Physiology and Biophysics, Faculty of Medicine, University of Calgary, Calgary, AB, Canada T2N 4N1.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: AC, auditory cortex; AS_r, repetitive acoustic stimuli; AS_t, train of acoustic stimuli; BF, best frequency; CF, constant frequency; CM, cochlear microphonic response; DSCF, Doppler-shifted constant frequency; ES_{ar}, repetitive electric stimulation of the AC; ES_l, electric stimulation of the leg; FM, frequency modulation; IC, inferior colliculus; MGB, medial geniculate body; SPL, sound pressure level.

detection (facilitation), coincidence detection associated with delay lines, and disinhibition (1, 5–7, 15, 21–24). All investigators have assumed that the mechanisms are caused by divergent and convergent projections of neurons in the ascending auditory system. Because the neurons with the response properties listed above have been found in the subcortical auditory nuclei, they are expected to be under corticofugal modulation. Recent findings indicate that the corticofugal system plays important roles in shaping (or even creating) the response properties of central auditory neurons and in reorganizing the cochleotopic and computational maps.

The Corticofugal Auditory System: Anatomy

Neurons in the deep layers of the AC project to the medial geniculate body (MGB), inferior colliculus (IC), or subcollicular auditory nuclei (25–27). These corticofugal projections are tonotopically organized (27, 28). Corticothalamic fibers project only to the ipsilateral MGB and thalamic reticular nucleus (27, 29). However, corticocollicular fibers bilaterally project to the IC. The ipsilateral projection is much more extensive and topographically organized than the contralateral projection (26). Therefore, ipsilateral corticofugal modulation is expected to be much larger than contralateral corticofugal modulation in the IC and MGB and to be frequency-dependent. Corticofugal projections are bilateral to the subcollicular nuclei: superior olivary complex and cochlear nucleus (30). Corticofugal modulation is expected to take place even in the cochlea via olivocochlear neurons in the superior olivary complex. The central nucleus of the IC projects not only to the MGB and the superior colliculus, but also to medial olivocochlear neurons, which mostly project to contralateral cochlear outer hair cells. In general, olivocochlear neurons bilaterally project to the cochlea, although there are some differences in olivocochlear projections between species (see ref. 31 for review).

Because the corticofugal system forms multiple feedback loops, the exploration of corticofugal functions is ongoing at different levels of the auditory system. An obvious critical experiment to be performed is the selective inactivation of individual feedback loops without injuring the ascending auditory system. Such an experiment, however, appears to be impossible because of anatomical complexity.

Corticofugal Modulation of Auditory Signal Processing

Gain Control. Physiological data of corticofugal effects on MGB and IC neurons have been controversial: (i) only or predominantly inhibitory (32–38); (ii) only or predominantly excitatory or facilitative (39–41); or (iii) equally excitatory or inhibitory (42, 43). These data, regardless of the excitatory or inhibitory effect, indicate that one of the corticofugal functions can be nonspecific gain control. In the mustached bat, nonfocal inactivation of cortical auditory neurons, including neurons matched to recorded subcortical neurons, evokes a large reduction of the auditory responses of the subcortical neurons (44, 45). Matched means that electrically stimulated cortical neurons and recorded subcortical or cortical neurons are tuned to the same value of an acoustic parameter. Unmatched means that they are tuned to different values of an acoustic parameter. One of the corticofugal functions is amplification of the responses of subcortical neurons. Because there is a much larger number of corticofugal fibers than thalamocortical fibers, the corticofugal system should have much more elegant functions than simple gain control.

To study the functional roles of the corticofugal system, one should not ignore that corticofugal and subcortical neurons both are tuned to particular values of an acoustic parameter. Therefore, electrical stimulation or drug application for activation or inactivation should be highly focal except for the initial phase of corticofugal research, and corticofugal effects on subcortical neurons should be evaluated with regard to the relationship in tuning between stimulated or inactivated cortical neurons and recorded subcortical neurons. The recent research designed on this philosophy leads us to several findings in the mustached bat, *Pteronotus parnellii parnellii*, and the big brown bat, *Eptesicus fuscus*.

Egocentric Selection in the Mustached Bat. In the mustached bat, DSCF neurons are extremely sharply tuned to frequencies at ≈ 61 kHz. They are specialized for processing frequency and/or amplitude modulated insect echoes. On the other hand, FM-FM neurons are combination-sensitive and are tuned to particular values of echo delays. They are specialized for processing target-distance information (see refs. 15 and 46 for review). To examine corticofugal modulation of the responses of thalamic and/or collicular DSCF or FM-FM neurons, the cortical DSCF (47, 48) or FM-FM area (49) was focally inactivated with 90 nl of 1.0% lidocaine or focally and repetitively activated with 100-nA, 0.2-ms long electric pulses delivered at a rate of 5/s for 7 min or 6.7/s for 6.7 min (ES_{ar} , repetitive electric stimulation of the AC).

DSCF Neurons. Cortical DSCF neurons, via the corticofugal system, mediate a highly focused positive feedback to augment the auditory responses at the BFs of matched thalamic or collicular DSCF neurons (hereafter, subcortical neurons). The BFs of these matched neurons are the same within ± 0.2 kHz as the BF of the electrically stimulated cortical neurons (hereafter, cortical BF). The positive feedback is very strong. Without it, the responses of subcortical neurons would be small. The positive feedback always is accompanied by widespread lateral inhibition, which suppresses the auditory responses at the BFs of unmatched subcortical DSCF neurons (hereafter, subcortical BF). The BFs of these subcortical neurons are different by more than 0.2 kHz, but not more than 2.0 kHz, from the cortical BF. Focal cortical inactivation with lidocaine evokes subcortical changes, which are exactly opposite to the above (Fig. 1 E and F).

Focal activation of cortical DSCF neurons modulates response magnitude of subcortical DSCF neurons, sharpens their frequency-tuning curves, and shifts the BFs and tuning curves of unmatched subcortical DSCF neurons away from the cortical BF. This centrifugal BF shift lasts up to 3 h, i.e., “recovers” in 3 h. Therefore, cortical neurons, via the corticofugal system, adjust and improve auditory information processing in the subcortical auditory nuclei. In other words, cortical neurons adjust and improve their own input. These corticofugal functions were named egocentric selection (49). The effects of egocentric selection are nearly two times larger for thalamic DSCF neurons than for collicular DSCF neurons (Fig. 1 C-F). In other words, cortical neurons adjust and improve their own inputs through multiple corticofugal feedback loops. Focal inactivation of cortical DSCF neurons results in corticofugal changes that are just opposite to those evoked by focal cortical activation (Fig. 1 A-D).

FM-FM Neurons. When cortical FM-FM neurons are electrically stimulated, exactly the same corticofugal effects as above were found on collicular FM-FM neurons, which are tuned to echo FM components after an emitted FM component with specific time delays corresponding to specific target distances. The best (echo) delay to excite them differs between neurons. Focal electrical stimulation of cortical FM-FM neurons facilitates the responses of matched collicular FM-FM neurons to a pair of FM sounds and sharpens their delay-tuning curves without shifting their best delays. The duration of facilitative response caused by the corticofugal positive feedback appears to be adjusted in the MGB by GABAergic inhibitory neurons in the thalamic reticular nucleus feedback (ref. 50; see ref. 51 for review). The matched collicular FM-FM neurons have the best delays which are within ± 0.4 ms of the best delay of the electrically stimulated

cortical FM-FM neurons. On the other hand, the electrical stimulation of cortical FM-FM neurons suppresses the auditory responses at the best delays of unmatched collicular FM-FM neurons and sharpens and shifts their best delays away from the best delays of electrically stimulated cortical neurons. This centrifugal shift lasts up to 3 h. The unmatched collicular FM-FM neurons have best delays that are different by more than 0.4 ms from the best delay of the electrically stimulated cortical FM-FM neurons (49).

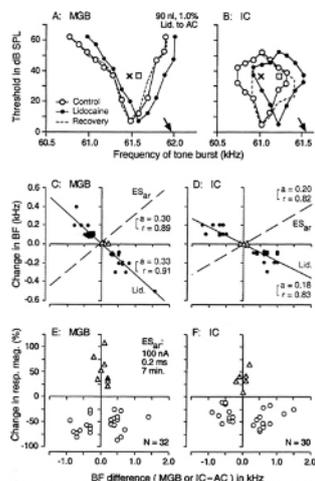


Fig. 1. Corticofugal modulation of thalamic (MGB) and collicular (IC) neurons in the mustached bat. Changes in the frequency-tuning curves of thalamic (A) and collicular neurons (B) by a focal inactivation of cortical neurons with 90 nl of 0.1% lidocaine (Lid-). The BFs of the inactivated cortical neurons are indicated by the arrows. The curves were measured before (control; \cdot), during (\circ), and after (recovery; dashed lines) the cortical inactivation. The tuning curves and BFs shift toward the BFs of inactivated cortical neurons. Crosses and squares indicate the best amplitudes measured before and during the cortical inactivation, respectively. Changes in the BFs of thalamic (C) and collicular neurons (D), i.e., reorganization of the frequency map, evoked by a focal inactivation of cortical neurons with lidocaine. The abscissae represent the differences in BF between cortical (AC) and thalamic (MGB) or collicular (IC) neurons in the control condition. The cortical BF was ≈ 61.2 kHz. The triangles and filled circles represent the data obtained from matched and unmatched subcortical neurons, respectively. The regression lines (solid), their slopes (a), and correlation coefficients (r) are shown. The dashed lines in C and D, respectively, show the regression lines for the BF shifts of thalamic and collicular neurons evoked by a focal cortical activation with 0.2-ms, 100-nA electric pulses at a rate of 5/s for 7 min (ES_{ar}). The slopes and correlation coefficients of these dashed lines also are shown. The BF shift was symmetrical and centrifugal for ES_{ar} . (E and F) The abscissae are the same as those in C and D. The ordinates represent percent change in the response magnitude (number of impulses per tone burst) of thalamic (E) and collicular (F) neurons evoked by ES_{ar} . The triangles and circles, respectively represent percent changes in response magnitude of matched and unmatched subcortical neurons at the BFs of individual neurons in the control condition. To measure response magnitudes, tone bursts were set at the best amplitude of each neuron in the control condition. Changes in BF (C and D) and response magnitude (E and F) both are larger in the MGB than in the IC (47, 48).

The corticofugal effects on subcortical DSCF and FM-FM neurons last long, so that it is hypothesized that egocentric selection is involved in the reorganization (plasticity) of the frequency and echo-delay maps of the central auditory system.

Corticofugal effects caused by positive feedback are stronger on subcortical FM-FM neurons than on subcortical DSCF neurons. Therefore, the processing of complex sounds by combination-sensitive neurons generally may depend on the corticofugal system more than does the processing by neurons primarily responding to single tones (44, 45).

Cochlear Hair Cells. The corticofugal system probably modulates the activity of cochlear hair cells through inhibitory olivocochlear neurons (26, 27, 31). It has been proposed that olivocochlear fibers improve the discrimination of complex sounds (52), increase the signal-to-noise ratio (53), increase the dynamic range of intensity coding (54), mediate selective attention (55), control the gain (56), and reduce temporary threshold shift (57). The functional role of the corticofugal system in signal processing at the subcollicular nuclei and the cochlea remains to be explored.

In the mustached bat, cochlear microphonic response (CM) recorded at the perilymphatic duct is sharply tuned to ≈ 61 kHz. Because of the sharp tuning, the CM evoked by a tone burst at ≈ 61 kHz shows a prominent off-response (CM-after), i.e., damped oscillations that occur at a fixed frequency (≈ 61 kHz) irrespective of the frequency of a stimulus tone burst (58). The focal electrical activation of the IC increases the resonance frequency and the duration of CM-after at ≈ 61 kHz (Fig. 2). Although the increase in the duration of CM-after suggests sharpening of the frequency tuning of outer hair cells, we don't yet know how our preliminary data are related to the modulation of auditory signal processing and whether focal activation of the cortical DSCF area representing ≈ 61 kHz evokes the same changes as those evoked by the focal activation of the IC.

Egocentric Selection in the Big Brown Bat. In the big brown bat, Jen *et al.* (38) found that electrical stimulation of the AC evoked either short latency facilitation (26%) or inhibition (74%) of collicular neurons regardless of whether they were matched or unmatched in BF with stimulated cortical neurons. They also found that the electric stimulation either augmented the auditory responses and broadened the frequency and spatial tuning curves of collicular neurons or suppressed the auditory responses and sharpened these tuning curves. The data obtained from the big brown bat by Suga and his coworkers (59–63), summarized below, are different from those obtained by Jen *et al.* (38), who used 0.1-ms long, 1.3- to 85- μ A pulses for cortical stimulation. Suga and his coworkers used 0.2-ms long, 0.1- μ A electric pulses for cortical stimulation.

Cortical neurons, via the corticofugal system, mediate focused positive feedback to augment the auditory responses at the BFs of matched collicular neurons without shifting their BFs and frequency tuning curves, as found in the mustached bat (Fig. 3 C and D). The BFs of matched neurons are within ± 0.5 kHz of the cortical BF.

The corticofugal effects on unmatched collicular neurons are inhibitory at their BFs and at frequencies higher than their BFs. However, the effects are facilitatory at the frequencies between the cortical and collicular BFs. The inhibitory and facilitative effects are larger on the high-frequency side than on the low-frequency side of the cortical BF (Fig. 3D). Because of these frequency-dependent corticofugal effects, the BFs and frequency-tuning curves of unmatched collicular neurons shift toward the cortical BF. This centripetal BF shift occurs predominantly for collicular neurons with BFs higher than the

cortical BF. Therefore, the centripetal BF shift is asymmetrical (Fig. 3C d and e). Asymmetrical and centripetal BF shifts also can be evoked by acoustic stimuli [20-ms long, 50-dB sound pressure level (SPL) tone bursts delivered at a rate of 10/s for 30 min; Fig. 3Cb].

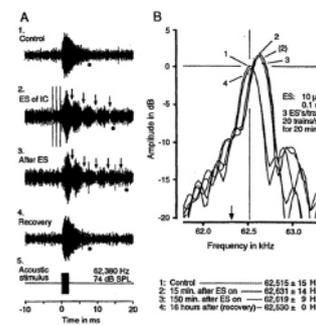


Fig. 2. Colliculofugal modulation of cochlear hair cells in the mustached bat. Changes in microphonic response (CM) to a 2.0-ms tone pulse evoked by electrical stimulation (ES) of the contralateral collicular neurons tuned to ≈ 62.32 kHz (arrows in *B*). (*A*) Oscillograms of CMs. (*B*) Amplitude spectra of the after-potentials of CMs (CM-after) shown in *A*. 1–4 were, respectively recorded before (control condition), during (15 min after the beginning of ES), 150 min after, and 16 h after ES (recovery condition). In *A*, the dots and arrows indicate the end of CM-after and nodes of beat, respectively. The lengthening of CM-after by ES suggests sharpening of frequency tuning of hair cells. In *B*, the peaks of the amplitude spectra indicate the resonance frequencies, which shifted from 62,515 Hz to 62,631 Hz (see the list at the bottom). The acoustic stimulus shown in *A*5 was a 2.0-ms, 62,380-Hz tone pulse at 74 dB SPL. The parameters of ES are listed in *B*. The three vertical lines in *A*2 are stimulus artifacts. The CMs were recorded from the cochlear perilymphatic duct with a tungsten-wire electrode.

BF shifts mean the reorganization of the cochleotopic (frequency) map of the IC. This reorganization can be easily demonstrated by dorso-ventral penetrations of a recording electrode along the frequency axis of the IC. In such a penetration, BFs systematically become higher with the electrode depth (Fig. 3A and B, \backslash). When cortical neurons with a particular BF (arrow in Fig. 3A or B) are electrically stimulated, the BF-depth curve shifts toward the cortical BF. This shift occurs for collicular BFs, which are 0–12 kHz higher than the cortical BF (Fig. 3A and B, \backslash). This reorganization of the frequency map results in the over-representation of the frequency equal to the cortical BF and the under-representation of frequencies that are 6–12 kHz higher than the cortical BF. Therefore, the contrast of the neural representation of the frequency of an acoustic stimulus is increased.

ES_{ar} also evokes BF shifts of cortical neurons located near the stimulation site that are very similar to collicular BF shifts. For example, centripetal BF shifts occur over 600–700 μ m rostral to the electrically stimulated cortical neurons with a 30-kHz BF and over 500 μ m caudal to these. The amount of BF shifts is asymmetrical and 3–4 times larger on the rostral (higher frequency) side of the stimulated cortical neurons than on the caudal (low frequency) side (62).

For 30-min ES_{ar}, collicular and cortical BF shifts develop up to $\approx 64\%$ of the plateau within 2 min, reach a plateau at 30 min, and then recover ≈ 180 min after the cessation of ES_{ar}. The recovery of BF shift tends to be slightly slower in the AC than in the IC (see Fig. 5A and B). The lengthening of ES_{ar} beyond 30 min, e.g., to 90 min, hardly increases the amount of BF shifts, but increases the duration of the plateau. The recovery time is nearly the same as that for the 30-min ES_{ar}. When the duration of ES_{ar} is shorter than 30 min, the BF shifts are small and recover quickly. For a 2-min ES_{ar}, the BF shifts are $\approx 64\%$ of the plateau and recover in ≈ 42 min. BF shifts in the big brown bat are associated with sharpening of frequency-tuning curves of some neurons (62), as found in the mustached bat (47).

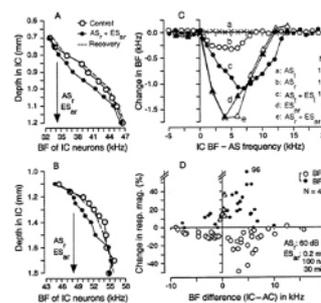


Fig. 3. Corticofugal modulation of collicular (IC) neurons in the big brown bat. (*A* and *B*) Shifts in the BFs of collicular neurons (i.e., reorganization of the frequency map) were evoked by a 100-nA, 0.2-ms electrical stimulation of cortical auditory neurons (ES_{ar}) paired with a 20-ms, 60-dB SPL acoustic stimulus (AS_r). AS_r+ES_{ar} was delivered at a rate of 10/s for 30 min. BF depth curves were obtained in two dorso-ventral electrode penetrations (*A* and *B*) across the IC 30 min before (control; \backslash), 30 min after (\backslash), and 2–3 h after (recovery; dashed lines) AS_r+ES_{ar}. Because almost all data points obtained 2–3 h after AS_r+ES_{ar} were the same as those obtained before the stimulation, they were expressed as a dashed curve without the individual data. The BFs of cortical neurons that were electrically stimulated are indicated by the arrows. (*C*) The amount of collicular BF shifts as a function of the difference between the BFs of collicular neurons and the frequency of acoustic stimulus (AS_r or AS_i) or the BF of electrically stimulated cortical neurons (ES_{ar}). AS_r was a 20-ms, 50-dB SPL tone burst delivered at a rate of 10/s for 30 min. AS_i was a 1.0-s train of tone bursts (10 ms each, 50 dB SPL, 33/s) delivered every 30 s for 30 min. AS_i was delivered alone or followed by electric stimulation (50 ms, 0.15–0.57 mA) of the leg (ES_i) with a 1.0-s gap as in trace conditioning. Each symbol indicates a mean of the data obtained from several electrode penetrations (*N*) across the IC. The standard error for each data point is not shown for simplicity. AS_i alone evoked no BF shift (*a*), but AS_r alone (*b*), AS_r+ES_i (*c*) or ES_{ar} alone (*d*) did. AS_r+ES_{ar} evoked BF shifts that were similar to those evoked by ES_{ar} alone (*e*). (*D*) Changes in response magnitude (number of impulses per tone burst) of collicular neurons evoked by AS_r+ES_{ar} are plotted as a function of the difference in BF between recorded collicular and electrically stimulated cortical neurons. The frequency of AS_r was the same as the BF of the electrically stimulated cortical neurons. The triangles and circles (open and filled) represent the data obtained from matched and unmatched neurons, respectively. The open and filled circles, respectively represent the changes in response magnitude at the BFs in the control condition (BF_C) and in the shifted condition (BF_S). The data were obtained with stimulus tone bursts at 10 or 20 dB above minimum threshold of each neuron. The frequencies of AS_i and AS_r for curves *a*–*c* in *C* were 25.3 ± 7.84 kHz. The BFs of cortical neurons stimulated by ES_{ar} for curves *d* and *e* in *C* were 39.5 ± 9.57 kHz (59, 60).

Corticofugal lateral inhibition of collicular or thalamic neurons may be based on (*i*) intrinsic cortical inhibition, which may adjust the amount of Corticofugal positive feedback, (*ii*) intrinsic thalamic inhibition, (*iii*) inhibition by the thalamic reticular

nucleus, and/or (iv) intrinsic collicular inhibition. Jen *et al.* (38) found that neurons in the external nucleus of the IC are excited by corticofugal fibers and in turn inhibit neurons in the central nucleus of the IC.

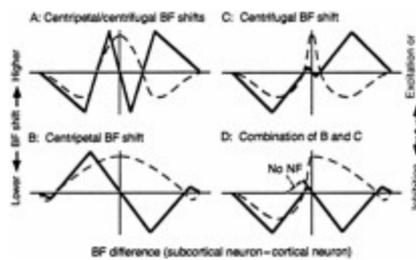


Fig. 4. The hypothesis to explain centripetal (solid red lines) and centrifugal BF shifts (solid blue lines) of subcortical neurons evoked by focal electrical stimulation of the auditory cortex. BF shifts (ordinates) are plotted as a function of difference in BF between recorded subcortical and activated cortical neurons (abscissae). The directions of BF shifts are hypothesized to depend on corticofugal excitation or facilitation (dashed red lines) and inhibition (dashed blue lines). NF, negative feedback. See the text.

Differences in Corticofugal Modulation Between Species and Between Ordinary and Specialized Areas

Egocentric selection has been found not only in the mustached bat (Fig. 1) and big brown bat (Fig. 3), but also in the cat (64), so that it may be a general function of the corticofugal system. Corticofugal positive feedback associated with lateral inhibition also has been found in the visual system (65). However, the effect of egocentric selection on the cochleotopic (frequency) map is different between different species of mammals and between ordinary and specialized areas of the AC of a single species.

Egocentric selection evokes centrifugal and symmetrical shifts of the BFs of DSCF neurons (Fig. 1) and the best delays of FM-FM neurons (49) of the mustached bat, but centripetal and asymmetrical BF shifts in the AC of the big brown bat (Fig. 3). BF shifts are centripetal and asymmetrical in the AC of the Mongolian gerbil (*Meriones unguiculatus*) and are centripetal and somewhat symmetrical in the posterior division of the AC of the mustached bat (66). The range and amount of the BFs shifted by ES_{ar} are also different from species to species and between the ordinary and specialized areas of the AC.

Corticofugal excitation or facilitation and inhibition appear to evoke centripetal and centrifugal BF shifts of unmatched subcortical neurons, respectively (Fig. 4A). If the excitatory effect is strong and widespread to neighboring unmatched neurons and negative feedback is weak, it may evoke prominent centripetal BF shifts (Fig. 4B). On the contrary, if the excitatory effect is highly focused to matched neurons and the inhibitory effect is strong and widespread to neighboring unmatched neurons, it may evoke prominent centrifugal BF shift (Fig. 4C). The corticofugal excitatory and inhibitory effects shown in Fig. 4 B and C may be combined in different ways (Fig. 4D). For example, if the excitatory effect is strong and widespread on the high-frequency side and there is neither excitatory nor inhibitory effect on the low-frequency side, it may evoke asymmetrical centripetal BF shift (Fig. 4D Right).

Neural representation of auditory signals in an ordinary AC appears to be improved by centripetal BF shifts, which result in over-representation of a particular value of a parameter characterizing a given acoustic stimulus. The area for over-representation is always bordered with the area or areas for under-representation. On the other hand, a specialized AC such as the DSCF and FM-FM areas over-represent particular values of a parameter in a narrow range in the natural condition. Therefore, further improvement for signal processing is performed by increasing the contrast in neural representation by centrifugal shifts in tuning curves. Asymmetrical centripetal BF shifts appear to be related to the asymmetrical shape of a frequency-tuning curve, as previously discussed (59). Regardless of types of BF shifts, tuning curves of neurons are sharpened by egocentric selection (47, 62).

Corticofugal Modulation and Plasticity of the Auditory System: Physiology and Behavior

The response properties of neurons and the sensory maps in a sensory cortex and subcortical sensory nuclei can be changed by conditioning, learning of a discrimination task, or focal cortical electrical stimulation (see refs. 67–69 for review). The importance of the corticofugal system to evoke plasticity in the central sensory system had not been considered until very recently (in the auditory system, refs. 47, 49, 51, and 59; in the somatosensory system, refs. 70 and 71). Gao and Suga (60, 61) exposed the big brown bat to various stimuli and obtained data indicating the importance of the corticofugal system in cortical plasticity. Their data and conclusions are summarized below.

Collicular and cortical neurons show asymmetrical and centripetal BF shifts not only for focal electrical stimulation of the AC (Fig. 3Cd), but also for a repetitive delivery of 20-ms long, 50-dB SPL tone bursts at a rate of 10/s for 30 min (AS_r , repetitive acoustic stimuli) (Fig. 3Cb). These findings indicate that egocentric selection is an intrinsic mechanism for the reorganization of the central auditory system.

A 1.0-s long train of 10-ms long, 50-dB SPL tone bursts at a burst rate of 33/s (AS_r , train of acoustic stimuli), delivered alone to the animal every 30 s for 30 min, does not evoke BF shift (Fig. 3Ca). However, when AS_r is delivered as a conditioned stimulus followed by an unconditioned electric leg stimulation (ES_l), large BF shifts are evoked that are also asymmetrical and centripetal (Fig. 3Cc). Neither ES_l alone nor AS_r delivered after ES_l (backward conditioning) evokes BF shifts. Because AS_r alone is behaviorally irrelevant, the data indicate that when an acoustic stimulus becomes behaviorally relevant to the animal, it evokes, via egocentric selection, considerable plasticity in the central auditory system, and that behavioral relevance is determined by the auditory and nonauditory systems, through associative learning.

The AC and the somatosensory cortex are both necessary for the BF shifts in the IC (60) and AC (61) caused by the conditioning, i.e., by associative learning. Electrical stimulation of the somatosensory cortex augments the collicular and cortical BF shifts evoked by the electric stimulation of the AC (63). Therefore, one of the nonauditory systems described above is the somatosensory cortex activated by the unconditioned leg stimulation. Another nonauditory system to be considered is the cholinergic basal forebrain, because its involvement in the plasticity of the AC has been demonstrated by Bakin and Weinberger (72) and Kilgard and Merzenich (73, 74).

Cortical and collicular BF shifts evoked by 30-min ES_{ar} are nearly the same in amount and recovery time (Fig. 5A a and b). However, those evoked by a 30-min conditioning session were quite different from each other (Fig. 5A c and d). Namely, the collicular BF shift is largest at the end of conditioning, and it is larger than the cortical BF shift within 45 min after the conditioning. The collicular BF shift recovers 180 min after the conditioning, just like that evoked by ES_{ar} (compare c with a in Fig. 5A). On the other hand, the cortical BF shift gradually increases after the conditioning, reaches a plateau at the time when the collicular BF shift almost recovers, and then stays over many hours (Fig. 5Ad). This is quite different from the cortical BF shift evoked by ES_{ar} (Fig. 5Ab). When the second condi

tioning session is given to the animal after the recovery of the collicular BF shift, it evokes the collicular BF shift, which is almost the same as that evoked by the first, and the cortical BF shift, which gradually increases over 3 h to a new plateau. However, the second conditioning session given to the animal at the beginning of the recovery phase of the collicular BF shift hardly changes the cortical BF shift, but prolongs the collicular BF shift. These observations indicate: (i) the collicular BF shift is not at all a consequence of the cortical BF shift, (ii) it always precedes the cortical BF shift, and (iii) its increasing phase is more related to the large cortical BF shift than its decreasing phase. The collicular BF shift appears to boost the cortical BF shift (61). Hereafter, we call plasticity lasting up to 3 h short term and plasticity lasting longer than 3 h long term.

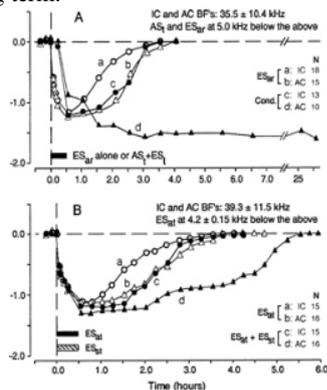


Fig. 5. Difference in time course of BF shift between collicular (IC) and cortical (AC) neurons evoked by focal cortical electrical stimulation or auditory conditioning in the big brown bat. (A) BF shifts were evoked by repetitive electric stimulation of the AC (ES_{ar}) (a and b) or the conditioning (cond.) (c and d). The conditioning consisted of a conditioned 1-s train of tone bursts (AS_t) followed by an unconditioned electric leg stimulus (ES_t). For ES_{ar} , cortical neurons showed a slightly slower BF recovery than did collicular neurons (a vs. b). For the conditioning, the BFs of cortical neurons slowly changed and did not recover within 1 day, but those of collicular neurons changed quickly and recovered as fast as those evoked by ES_{ar} (c vs. d) (61). (B) Collicular (a) and cortical BF shifts (b) evoked by trains of electric stimuli delivered to the AC (ES_{at}) were augmented mainly in duration by electrical stimulation of the somatosensory cortex (ES_{st}), as shown by c and d. ES_{at} and ES_{st} were delivered to mimic the conditioned (AS_t) and unconditioned stimuli (ES_t) (63). The mean BF of collicular and cortical neurons recorded was 35.5 ± 10.4 kHz for A and 39.3 ± 11.5 kHz for B. The frequencies of AS_t were always 5.0 kHz lower than the recorded collicular and cortical BFs. The BFs of cortical neurons stimulated by ES_{ar} or ES_{at} were 4.2 ± 0.15 kHz lower than the BFs of the recorded neurons. BF shifts were measured with tone bursts at 10 dB above minimum threshold of individual collicular or cortical neurons. N, number of neurons studied.

Stimulation of the basal forebrain with 0.2-ms, 100- μ A electric pulses immediately before and/or during (but not after) ES_{ar} augments the collicular and cortical BF shifts evoked by ES_{ar} , i.e., by the AC and the corticofugal system. The cortical BF shift becomes long-lasting, but the lengthening the recovery time of the collicular BF shift is small (63).

By 1990, a number of important findings on learning and memory had been made: (i) Acetylcholine plays an important role in learning and memory (see ref. 75 for review), (ii) The cholinergic basal nucleus of the forebrain projects diffusely and widely to the cerebral cortex (see ref. 76 for review), (iii) The basal forebrain plays an important role in learning and memory (77–79), (iv) The basal forebrain receives an input from the amygdala, which is necessary for the acquisition of conditioned response (80–82), (v) The amygdala receives an input from thalamic nuclei (83). Weinberger (see refs. 69 and 83 for review) pointed out the importance of the amygdala and cholinergic basal forebrain for plasticity of the AC evoked by fear conditioning. The basal forebrain undoubtedly plays an important role in the plasticity of the AC (72–74).

In the big brown bat, an acetylcholine application to the AC during auditory conditioning augments both collicular and cortical BF shifts, which were barely evoked by a 15-min conditioning session. An atropine application to the AC during the 30-min conditioning session completely abolishes the cortical BF shift and reduces the collicular BF shift (84). These observations indicate that the cholinergic system can augment the plasticity evoked by the AC and the corticofugal system and that the collicular BF shift can be evoked without the cortical BF shift. These observations suggest that the subcortical short-term change caused by egocentric selection, with the help of acetylcholine, boosts the cortical change into a long-term change.

Working Hypothesis

We propose the following working hypothesis of cortical plasticity. When behaviorally irrelevant acoustic stimuli are delivered to an animal, auditory signals representing the stimuli ascend from the cochlea to the AC. Then, the AC and the corticofugal system perform egocentric selection, which is a small and short-term modulation of subcortical signal processing. Accordingly, the small and short-term cortical change is evoked (Fig. 6 Left). When the acoustic stimuli are paired with electric leg stimulation, the auditory and somatosensory signals ascend from the periphery to the auditory and somatosensory cortices, respectively (Fig. 6 Center), and then to the amygdala through association cortices. These signals are associated in the amygdala, which is essential for evoking conditioned behavioral response. Therefore, the acoustic stimuli become behaviorally relevant to the animal. The amygdala sends the “associated” signal to the cholinergic basal forebrain, which increases the

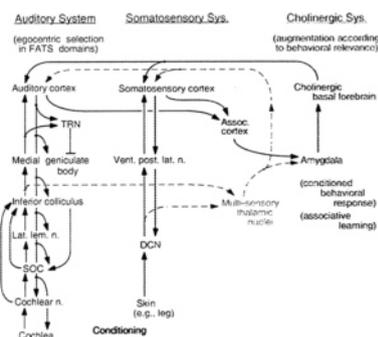


Fig. 6. Block diagram to explain our working hypothesis for the adjustment and improvement of auditory signal processing according to associative learning. DCN, dorsal column nuclei in the spinal cord. FATS, frequency, amplitude, time, and space. SOC, superior olivary complex. TRN, thalamic reticular nucleus. See the text.

cortical acetylcholine level (Fig. 6 Right). Then, the change in the AC is augmented. Accordingly, egocentric selection is augmented, and the subcortical change becomes larger, so that the cortical change becomes larger and long term. This positive feedback loop is controlled by inhibition mediated by the thalamic reticular nucleus.

The dashed arrows in Fig. 6 indicate pathways through the multisensory thalamic nuclei that have been considered to be essential for cortical plasticity by Weinberger (see refs. 69 and 83 for review). These pathways appeared to be quite reasonable to us. However, the importance of these pathways is doubtful because of the following data obtained from the big brown bat (61). (i) Inactivation or activation of the somatosensory cortex, respectively abolishes or augments cortical and collicular BF shifts evoked by fear conditioning. (ii) Collicular changes evoked by the corticofugal system precede cortical changes. (iii) The AC and corticofugal system have an intrinsic mechanism for cortical and subcortical plastic changes, which are highly specific to the parameters characterizing an acoustic stimulus, (iv) Fear conditioning evokes plasticity in the IC, which is the nucleus one step below the thalamus. In other words, the multisensory thalamic nuclei are not the first place where the plasticity caused by the conditioning is evoked.

Multiple Functions of the Corticofugal System for Hearing

The corticofugal function consists of at least the following subfunctions.

Egocentric Selection for Short-Term Adjustment and Improvement of Auditory Signal Processing According to Auditory Experience. Egocentric selection is performed by the AC and the corticofugal system. It is based on focused positive feedback associated with lateral inhibition. Egocentric selection adjusts and improves cortical neurons' own input, so that it is similar to the function of corticofugal feedback in the visual system (85).

Reorganization for Long-Term Adjustment and Improvement of Signal Processing. Such reorganization is based on egocentric selection working together with nonauditory systems.

Gain Control. Overall facilitative or inhibitory corticofugal modulation indicate this function. Egocentric selection can be viewed as selective gain control.

Shaping (or Even Creating) Response Properties of Neurons, in Particular, of Combination-Sensitive Neurons. Filter properties of neurons in the frequency (37, 38, 47, 62), amplitude, time (49), and spatial domains (38) can be sharpened by the corticofugal system. Sharpening of tuning is one of the functions of egocentric selection. The creation of combination sensitivity through interactions between the ascending and descending systems (45) must be a particularly important function for processing behaviorally relevant complex sounds. In the visual system, response properties of thalamic neurons become complex because of corticofugal feedback (86).

Binding of the Different Features of Auditory Signals. The problem of binding has been extensively studied in the visual system, but not in the auditory system. In the visual system, the corticofugal system evokes feature-linked synchronized discharges in the thalamic neurons (86, 87).

Stabilization of Thalamic Auditory Responses Via the Thalamic Reticular Nucleus. The thalamic reticular nucleus receives axon collaterals from both ascending thalamo-cortical fibers and descending cortico-thalamic fibers. Corticofugal positive feedback has a high gain, so that ringing would be evoked if it is not incorporated with inhibition through the thalamic reticular nucleus (see ref. 51 for review). If the thalamic reticular nucleus does not operate properly, long-lasting discharges, perhaps responsible for tinnitus, would be produced. Cooling of the AC had "complex" effects on the auditory responses of MGB and reticular nuclear neurons, so that Villa *et al.* (41) proposed that the reticular nucleus takes a role as an adaptive filter.

Attentional Modulation of Auditory Signal Processing. In cats, visual attention to a mouse reduces auditory responses of the dorsal cochlear nucleus (88) and a visual discrimination task reduces auditory nerve responses to clicks (89). In humans, visual attention reduces auditory nerve responses (90) and otoacoustic emissions evoked by a click (91). The corticofugal system probably mediates attentional modulation of auditory signal processing.

Low-Frequency Modulation of Brain Rhythm. The corticofugal system transmits slow oscillatory changes in cortical activity to the thalamic visual nucleus. This slow oscillation (0.6–1.0 Hz) interacts with spindles (7–14 Hz) generated in the thalamus, modulates neural excitability, and produces different brain rhythms characterizing various behavioral states (see ref. 92 for review).

We thank Drs. S.P. Dear, K.K. Ohlemiller, J.J. Wenstrup, and Mr. N. Laleman for their comments on this paper. This work has been supported by a research grant from the National Institute on Deafness and Other Communicative Disorders (DC 00175).

- Suga, N., Zhang, Y. & Yan, J. (1997) *J. Neurophysiol.* **77**, 2098–2114.
- Suga, N. & Schlegel, P. (1973) *J. Acoust. Soc. Am.* **54**, 174–190.
- Covey, E. & Casseday, J.H. (1991) *J. Neurosci.* **11**, 3456–3470.
- Sullivan, W.E. (1982) *J. Neurophysiol.* **48**, 1011–1132.
- Casseday, J.H., Ehrlich, D. & Covey E. (1994) *Science* **264**, 847–850.
- Suga, N. (1965) *J. Physiol. (London)* **181**, 671–700.
- Suga, N. (1969) *J. Physiol. (London)* **200**, 555–574.
- Fuzessery, Z.M. (1994) *J. Neurophysiol.* **72**, 1061–1079.
- Suga, N. & Manabe, T. (1982) *J. Neurophysiol.* **47**, 225–255.
- Haplea, S., Covey, E. & Casseday, J.H. (1994) *J. Comp. Physiol. A* **174**, 671–683.
- Casseday, J.H., Covey, E. & Grothe, B. (1997) *J. Neurophysiol.* **77**, 1595–1605.
- Suga, N., O'Neill, W.E., Kujirai, K. & Manabe, T. (1983) *J. Neurophysiol.* **49**, 1573–1626.
- Irvine, D.R.F. (1992) in *The Mammalian Auditory Pathway: Neurophysiology*, eds. Popper, A.N. & Fay, R.R. (Springer, New York), pp. 153–231.
- Galazyuk, A.V. & Feng, A.S. (1997) *J. Comp. Physiol. A* **180**, 301–311.
- Suga, N. (1994) in *The Cognitive Neurosciences*, ed. Gazzaniga, M.S. (MIT Press, Cambridge, MA), pp. 295–318.
- Valentine, D.E. & Moss, C.F. (1997) *J. Neurosci.* **17**, 1720–1733.
- Manabe, T., Suga, N. & Ostwald, J. (1978) *Science* **200**, 339–342.
- Imig, T.J. & Adrián, H.O. (1977) *Brain Res.* **138**, 241–257.
- Kujirai, K. & Suga, N. (1983) *Auris Nasus Larynx (Tokyo)* **10**, 9–24.
- Jen, P.H.-S., Sun, X.D. & Lin, P.J.J. (1989) *J. Comp. Physiol.* **165**, 1–14.
- Suga, N. & Tsuzuki, K. (1985) *J. Neurophysiol.* **53**, 1109–1145.
- Yang, L., Pollak, G.D. & Resler, C. (1992) *J. Neurophysiol.* **68**, 1760–1774.
- Fuzessery, Z.M. & Hall, J.C. (1996) *J. Neurophysiol.* **76**, 1059–1053.
- Jen, P.H.-S. & Feng, R.B. (1999) *J. Comp. Physiol. A* **184**, 185–194.
- Kelly, J.P. & Wong, D. (1981) *Brain Res.* **212**, 1–15.
- Saldana, E., Feliciano, M. & Mugnaini, E. (1996) *J. Comp. Neurol.* **371**, 15–40.
- Huffman, R.F. & Henson, O.W., Jr. (1990) *Brain Res. Rev.* **15**, 295–323.
- Malmierca, M.S., Le Beau, F.E.N. & Rees, A. (1996) *Hear. Res.* **93**, 167–180.
- Ojima, H. (1994) *Cereb. Cortex* **6**, 646–663.
- Feliciano, M., Saldana, E. & Mugnaini, E. (1995) *Aud. Neurosci.* **1**, 287–308.
- Warr, W.B. (1992) in *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D.B., Popper, A.N. & Fay, R.R. (Springer, New York), pp. 410–448.
- Desmedt, J.E. & Mechelse, K. (1958) *Proc. Soc. Exp. Biol.* **99**, 772–775.
- Massopust, L.C., Jr. & Ordly, J.M. (1962) *Exp. Neurol.* **6**, 465–477.
- Watanabe, T., Yanagisawa, K., Kanzaki, J. & Katsuki, Y. (1966) *Exp. Brain Res.* **2**, 302–317.
- Amato, G., La Grutta, V. & Enia, F. (1969) *Arch. Sci. Biol.* **53**, 291–313.
- Sun, X., Jen, P.H. S., Sun, D. & Zhang, S. (1989) *Brain Res.* **495**, 1–8.
- Sun, X., Chen, Q. & Jen, P.H. (1996) *Neurosci. Lett.* **212**, 131–134.

38. Jen, P.H.-S., Chen, Q.C. & Sun, X.D. (1998) *J. Comp. Physiol* **183**, 683–697.
39. Andersen, P., Junge, K. & Sveen, O. (1972) *Brain Behav. Evol.* **6**, 170–184.
40. Orman, S.S. & Humphrey, G.L. (1981) *Exp. Brain Res.* **42**, 475–482.
41. Villa, A.E., Rouiller, E.M., Simm, G.M., Zurita, P., de Ribaupierre, Y. & de Ribaupierre, F. (1991) *Exp. Brain Res.* **86**, 506–517.
42. Ryugo, D.K. & Weinberger, N.M. (1976) *Exp. Neurol.* **51**, 377–391.
43. Syka, J. & Popelar, J. (1984) *Neurosci. Lett.* **51**, 235–240.
44. Zhang, Y. & Suga, N. (1997) *J. Neurophysiol.* **78**, 3489–3492.
45. Yan, J. & Suga, N. (1999) *J. Neurophysiol.* **81**, 817–824.
46. Suga, N. (1990) *Sci. Am.* **262**, 60–68.
47. Zhang, Y., Suga, N. & Yan, J. (1997) *Nature (London)* **387**, 900–903.
48. Zhang, Y. & Suga, N. (2000) *J. Neurophysiol.* **84**, 325–333.
49. Yan, J. & Suga, N. (1996) *Science* **273**, 1100–1103.
50. Butman, J.A. (1992) Ph.D. thesis (Washington University, St. Louis).
51. Suga, N., Butman, J.A., Teng, H., Yan, J. & Olsen, J.F. (1995) in *Active Hearing*, eds. Flock, A., Ottoson, D. & Ulfendahl, M. (Elsevier, London), pp. 13–30.
52. Dewson, J.H. (1968) *J. Neurophysiol.* **31**, 122–130.
53. Kawase, T., Delgutte, B. & Liberman, M.C. (1993) *J. Neurophysiol.* **70**, 2533–2549.
54. Geisler, C.D. (1974) *J. Acoust. Soc. Am.* **56**, 1910–1902.
55. Oatman, L.C. & Anderson, B.W. (1977) *Exp. Neurol.* **57**, 200–211.
56. Siegel, J.H. & Kim, D.O. (1982) *Hear. Res.* **6**, 171–182.
57. Rajan, R. (1990) *Brain Res.* **506**, 192–204.
58. Suga, N. & Jen, P.H.-S. (1977) *J. Exp. Biol.* **69**, 207–232.
59. Yan, W. & Suga, N. (1998) *Nat. Neurosci.* **1**, 54–58.
60. Gao, E. & Suga, N. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 12663–12670.
61. Gao, E. & Suga, N. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 8081–8086.
62. Chowdhury, S. & Suga, N. (2000) *J. Neurophysiol.* **83**, 1856–1863.
63. Ma, X. & Suga, N. (2000) *Assoc. Res. Otolaryngol. Abstr.* p. 259, no. 899.
64. He, J.F. (1997) *J. Neurophysiol.* **77**, 896–908.
65. Tsumoto, T., Creutzfeldt, O.D. & Legendy, C.R. (1978) *Exp. Brain Res.* **32**, 345–364.
66. Sakai, M. & Suga, N. (2000) *Soc. Neurosci. Abstr.*, in press.
67. Irvine, D.R.F. & Rajan, R. (1996) *Clin. Exp. Pharmacol. Physiol.* **23**, 939–947.
68. Buonomano, D.V. & Merzenich, M.M. (1998) *Annu. Rev. Neurosci.* **21**, 149–186.
69. Weinberger, N.M. (1998) *Neurobiol. Learning Memory* **70**, 226–251.
70. Ergenzinger, E.R., Glasier, M.M., Hahm, J.O. & Pons, T.P. (1998) *Nat. Neurosci.* **1**, 226–229.
71. Krupa, D.J., Ghazanfar, A.A. & Nicolelis, M.A.L. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 8200–8205.
72. Bakin, J.S. & Weinberger, N.M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11219–11224.
73. Kilgard, M.P. & Merzenich, M.M. (1998) *Science* **279**, 1714–1718.
74. Kilgard, M.P. & Merzenich, M.M. (1998) *Nat. Neurosci.* **1**, 727–731.
75. Bartus, R.T., Dean, R.L., III, Beer, B. & Lipps, A.S. (1982) *Science* **217**, 408–414.
76. Johnston, M.V., McKinney, M. & Coyle, J.T. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5392–5396.
77. Bartus, R.T., Flicker, C., Dean, R.L., Pontecorvo, M., Figueiredo, J.C. & Fisher, S.K. (1985) *Pharmacol. Biochem. Behav.* **23**, 125–135.
78. Rigdon, G.C. & Pirch, J.H. (1986) *J. Neurosci.* **6**, 2535–2542.
79. Wozniak, D.F., Stewart, G.R., Finger, S., Olney, J.W. & Cozzari, C. (1989) *Neurobiol. Aging* **10**, 173–179.
80. Krettek, J.E. & Price, J.L. (1978) *J. Comp. Neurol.* **178**, 225–254.
81. Price, J.L. & Amaral, D.G. (1981) *J. Neurosci.* **1**, 1242–1259.
82. Peinado-Manzano, A. (1988) *Behav. Brain Res.* **29**, 61–71.
83. Weinberger, N.M. (1990) *Concepts Neurosci.* **1**, 91–123.
84. Ji, W., Suga, N. & Gao, E. (2000) *Soc. Neurosci. Abstr.*, in press.
85. Murphy, P.C., Duckett, S.G. & Sillito, A.M. (1999) *Science* **286**, 1552–1554.
86. Sillito, A.M., Jone, H.E., Gerstein, G.L. & West, D.C. (1994) *Nature (London)* **369**, 479–482.
87. Gray, C.M., König, P., Engel, A.K. & Singer, W. (1989) *Nature (London)* **338**, 334–337.
88. Hernandez-Peon, R., Scherrer, H. & Jouvet, M. (1956) *Science* **123**, 331–332.
89. Oatman, L.C. (1971) *Exp. Neurol.* **32**, 341–356.
90. Lukas, J.H. (1980) *Psychophysiology* **17**, 444–452.
91. Puel, J.L., Bonfils, P. & Pujol, R. (1988) *Brain Res.* **447**, 380–383.
92. Steriade, M. (1999) *Trends Neurosci.* **22**, 337–345.

TRACES OF LEARNING IN THE AUDITORY LOCALIZATION PATHWAY

Eric I. Knudsen*, Weimin Zheng, and William M. DeBello

Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305

One of the fascinating properties of the central nervous system is its ability to learn: the ability to alter its functional properties adaptively as a consequence of the interactions of an animal with the environment. The auditory localization pathway provides an opportunity to observe such adaptive changes and to study the cellular mechanisms that underlie them. The midbrain localization pathway creates a multimodal map of space that represents the nervous system's associations of auditory cues with locations in visual space. Various manipulations of auditory or visual experience, especially during early life, that change the relationship between auditory cues and locations in space lead to adaptive changes in auditory localization behavior and to corresponding changes in the functional and anatomical properties of this pathway. Traces of this early learning persist into adulthood, enabling adults to reacquire patterns of connectivity that were learned initially during the juvenile period.

The location of a sound source does not project directly onto the sensory surface of the ear. Therefore, the central auditory system must infer the location of a stimulus from acoustic cues created by the interaction of the incoming sound with the head and ears. Interaural time difference (ITD), which results from a difference in the distance that sound travels to reach the near versus far ear, is the dominant cue for the horizontal (azimuthal) position of a sound source (1, 2). Other cues, such as interaural level differences (ILDs) and monaural spectral cues, which result from the frequency-dependent directional properties of the head and ears, have more complicated relationships with the horizontal and vertical locations of a sound source (3, 4). These cues are used for localizing sounds in both azimuth and elevation (2, 5).

The central auditory system processes localization cues in parallel pathways in the midbrain and forebrain (6). The midbrain localization pathway (Fig. 1) branches from the main, tonotopic pathway at the level of the central nucleus of the inferior colliculus (ICC). This pathway leads to the optic tectum (called the superior colliculus in mammals), a primary function of which is to orient an animal's gaze toward interesting stimuli (9, 10). The forebrain localization pathway branches from the tonotopic pathway at the level of the primary auditory field in the forebrain (11). This pathway leads to various forebrain areas subserving a variety of behaviors, including auditory spatial memory and gaze control (12).

To guide behavior accurately, the midbrain and forebrain localization pathways must associate values of localization cues with appropriate locations in space. The establishment and maintenance of an accurate representation of space is complicated by variability in the correspondences of cues with locations across sound frequencies and across individuals (3, 13). Moreover, the encoding of cue values by nerve impulses will vary with changes in the relative sensitivities of the ears and with the development and aging of the nervous system. The midbrain and forebrain pathways deal with the variability in cue-location correspondences by adjusting cue-location associations based on experience (14–16). The structural and functional changes that underlie these experience-driven adjustments are the subject of this review.

The shaping influence of experience on the auditory localization pathway has been documented most thoroughly in the midbrain pathway of the barn owl, a nocturnal predator that depends on sound localization for survival. Therefore, this review focuses primarily on these data. It is likely, however, that the principles of adaptive plasticity that operate in the midbrain pathway of barn owls operate in other pathways and in other species as well.

Physiological Traces of Learning

Behavioral adjustments to manipulations of auditory or visual experience leave clear physiological traces in the midbrain localization pathway of owls and ferrets (17, 18). In the optic tectum, auditory space is represented topographically, and the tuning of neurons for sound localization cues is sharp. Many tectal neurons also respond to visual stimuli, and their auditory and visual spatial receptive fields are mutually aligned. The location of a tectal neuron's visual receptive field indicates with high precision the values of sound localization cues to which a neuron should be tuned normally and, therefore, can be used to detect changes in the tuning of tectal neurons that result from abnormal sensory experience (Fig. 2).

Abnormal Auditory Experience. Chronic occlusion of one ear in young barn owls leads to adaptive adjustments in auditory localization behavior and in auditory spatial tuning in the optic tectum (19). By changing the timing and level of sound reaching the eardrum, an earplug alters the relationship of ITDs and ILDs with locations in space. Initially, monaurally occluded owls mislocalize sounds toward the side of the unplugged ear. After experiencing an earplug for a period of months, however, owls recover accurate auditory orienting responses. In the optic tecta of these animals, neurons exhibit altered tuning for ITD and ILD so that their auditory receptive fields align with their visual receptive fields despite the presence of the earplug (20, 21). Analogous adaptive adjustments in ITD and ILD tuning can also result from experimental alterations of the directional properties of the external ears in young and adult owls (22).

Adaptive adjustments to abnormal auditory experience are carried out in a frequency-dependent manner in the midbrain pathway. Values of ITD and ILD can be altered independently

*To whom reprint requests should be addressed.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: ITD, interaural time difference; ILD, interaural level difference; ICC, the central nucleus of the inferior colliculus; ICX, the external nucleus of the inferior colliculus; NMDA, *N*-methyl-D-aspartate; AP5, *D,L*-2-amino-5-phosphonopentanoic acid; GABA, γ -aminobutyric acid.

for different frequencies of sound by installing an acoustic filtering device in one ear (23). Acute insertion of such a device shifts the locations of tectal auditory receptive fields in different directions and by different amounts, depending on the frequency of the sound used to measure the receptive fields. After young owls experience such a device for a period of months, auditory receptive fields measured with different frequencies of sound are realigned with each other and with visual receptive fields, as long as the device is in the ear. The basis for this adaptive adjustment is a frequency-specific change in the tuning of tectal neurons for ITD and ILD that reflects the acoustic effects of the device (24).

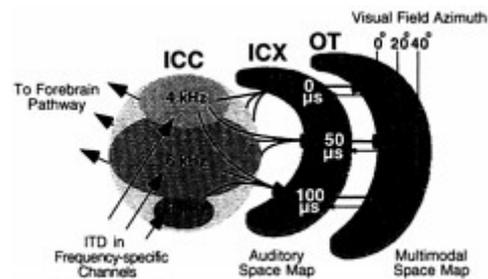


Fig. 1. The midbrain auditory localization pathway. For clarity, not all connections or pathways are shown. ITD is initially measured and mapped in frequency-specific channels in the nucleus laminaris, the avian equivalent of the mammalian medial superior olive (7). This information ascends to the ICC (arrows). From the ICC, information proceeds to the auditory thalamus in the forebrain and to the ICX in the midbrain pathway. In the projection from the ICC to the ICX, information about cue values, including ITD, converges across frequency channels, resulting in spatially restricted auditory receptive fields and a map of space. The auditory map of space merges with a visual map of space in the optic tectum. Tectal neurons receive visual inputs directly from the retina and indirectly from the forebrain. The optic tectum also projects topographically back to the ICX (8). OT, optic tectum.

Abnormal Visual Experience. Auditory orienting behavior and auditory spatial tuning in the optic tectum also are adjusted adaptively in response to optical displacements of the visual field (17). At first, it may seem surprising that auditory responses should shift as a result of a shift in the visual field. Recall, however, that a primary function of the optic tectum is to orient gaze toward interesting stimuli. Prisms do not affect the movement required to bring a visual stimulus (seen through the prisms) onto the center of gaze. In contrast, prisms do alter the movement required to bring an auditory stimulus onto the center of gaze. Realignment of auditory receptive fields with optically displaced visual receptive fields in the tectum results in orienting movements that enable the owl to foveate the source of the auditory stimulus through the prisms. Because the eyes of a barn owl do not move substantially relative to the head (movement $< 2^\circ$), auditory receptive fields must change to realign with optically displaced visual receptive fields.

Horizontally displacing prisms cause a horizontal shift of auditory receptive fields in the optic tecta of young owls (25). This shift in auditory spatial tuning is caused by a shift in the tuning of tectal neurons for ITD (Fig. 2), the dominant cue for azimuth. The dynamics of this shift in tuning can be observed with high resolution in the tectum (26). Learned, adaptive neuronal responses gradually increase over a period of weeks after mounting prisms. With additional experience, normal responses are gradually eliminated and the remaining responses become sharply tuned for learned values of ITD. Thus, a shift in ITD tuning involves two components: (i) acquisition of neuronal responses to values of ITD that are adaptive with the prisms in place (“learned responses”) and (ii) elimination of responses to values of ITD that previously were effective but are no longer appropriate with the prisms in place (“normal responses”).

Site of Plasticity. Abnormal auditory or visual experience leads to adaptive changes at the same site in the midbrain localization pathway: the external nucleus of the inferior colliculus (ICX). The ICX is the site in the pathway where the auditory system creates a map of space by merging information across frequency channels (Fig. 1). The frequency-specific information about cue

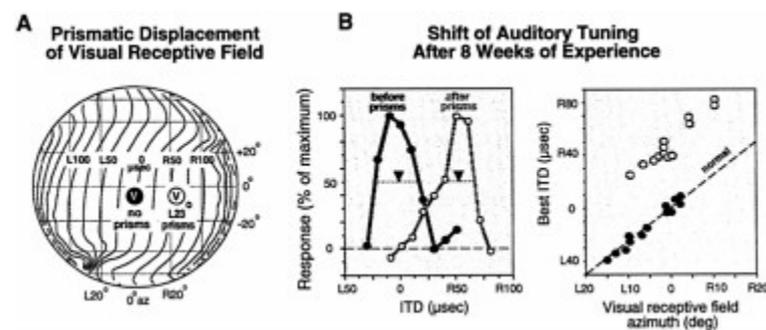


Fig. 2. Functional traces of learning in the optic tectum of the barn owl. Experience with a prismatically displaced visual field causes adaptive adjustment of ITD tuning in the auditory space map. (A) The effect of $L23^\circ$, horizontally displacing prisms on a visual receptive field location (encircled “V”). The globe represents space in front of the owl relative to the line of sight. The contour lines indicate the correspondence of ITD values (for 6 kHz sounds) with locations in space. (B) The ITD tuning of tectal neurons before and after prism rearing. The visual receptive fields of both neurons are centered at 0° azimuth. Before prism experience (blue symbols), the neuron is tuned to $0 \mu\text{s}$ ITD. After the owl has experienced $L23^\circ$ prisms for 8 weeks, the ITD tuning of the neuron from a very similar site has shifted to $50 \mu\text{s}$ right-ear leading (purple symbols). Downward arrows indicate the best ITD for each site. Best ITD is the midpoint of the range of ITDs that elicit more than 50% of the maximum response for that neuron. (C) Eight weeks of prism experience causes the relationship between best ITD and visual receptive field azimuth to be shifted (purple symbols) systematically from normal (blue symbols). The dashed line indicates the normal regression of best ITD on visual receptive field azimuth. The location of visual receptive fields, measured with prisms removed, is not altered in the tectum by prism experience. L, left-ear leading; R, right-ear leading. Data are from Brainard and Knudsen (29).

values is provided by inputs from neurons in the ICC. By combining information across frequency channels, ICX neurons eliminate much of the spatial ambiguity that is inherent in individual, frequency-specific cues (27, 28).

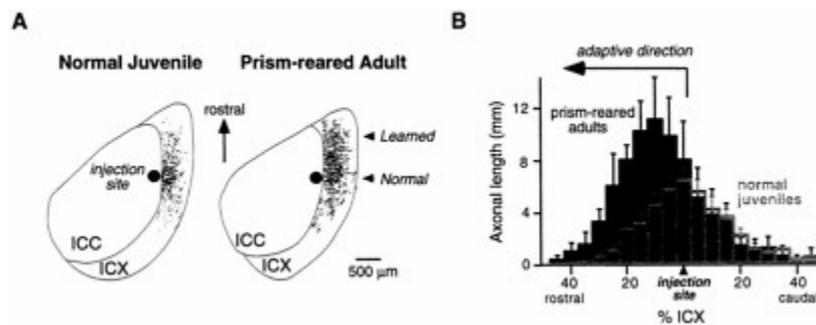


Fig. 3. Anatomical traces of learning in the ICX of the barn owl. (A) Digital image drawings of labeled axons in 40- μ m thick, horizontal sections through the optic lobe of a normal juvenile (*Left*) and a prism-reared adult (*Right*). To visualize the pattern of axonal projection from the ICC to the ICX, the anterograde tracer, biocytin, was injected into the ICC at a site representing 20 μ s ITD contralateral ear leading. The location of the injection site in the ICC is indicated as a dark circle. Labeled axons in the ICX are shown as thin lines. In a normal juvenile, the axonal projection field in the ICX is spatially restricted, symmetric, and centered around the rostral-caudal level of the injection site. In contrast, in a prism-reared adult expressing a rostral map shift, there is a dramatic expansion of the projection field in the rostral portion of the ICX. The direction of this axonal expansion is adaptive for the direction of prismatic displacement. Thus, these axons represent the learned projection field. (B) Composite spatial distributions of axons for normal juveniles (empty gray bars) and prism-reared adults (solid black bars). To quantify the spatial distribution for each case, the ICX was subdivided into measurement zones oriented orthogonal to the rostral-caudal axis. Each measurement zone corresponded to 5% of the rostral-caudal extent of the ICX. Total axonal length in each measurement zone was determined by computer analysis of the digital image drawings. Histograms were aligned relative to the rostral-caudal level of the injection site. Composite distributions were obtained by averaging the individual cases. In prism-reared adults with rostral map shifts ($n=4$), the axonal density on the rostral flank of the distribution is significantly greater than in normal juveniles ($n=7$; ANOVA, $P < 0.0001$), indicating that remodeling occurs by a net elaboration of axons. In contrast, the axonal density within the normal projection field, located at the rostral-caudal level of the injection site, is not significantly changed by prism experience. Data are from W.M.D., D.Feldman, and E.I.K., unpublished data.

In owls that have adjusted to either optical prisms or an acoustic filtering device, the adaptive functional changes in the auditory space map observed in the optic tectum also are observed in the ICX (16, 29). In contrast, the representations of cue values in the ICC, in both cases, remain unaltered. Normal representations in the ICC followed by adaptively shifted representations in the ICX suggest, but do not prove, that a site of experience-driven plasticity is the ICX. Proof that the ICX is a site of plasticity is based on the anatomical and pharmacological traces of learning discussed below.

That adaptive adjustments to abnormal auditory experience are made at the level of the ICX makes sense because the adjustments that need to be made are different for different frequency channels; therefore, the adjustments must be made before information is merged across frequency channels. Adaptive adjustments in response to optical prisms may occur at the same site simply because the nervous system normally uses vision to calibrate changing and unpredictable auditory cues and has not evolved to deal with large, sustained displacements of the visual field.

Anatomical Traces of Learning

Adaptive adjustments to prisms involve anatomical changes in the midbrain localization pathway (ref. 30 and W.M.D., D. Feldman, and E.I.K., unpublished data). Normally, neurons located in the lateral regions of the ICC send axons radially into the ICX (Fig. 3A *Left*). This topographic projection forms the basis of the space map (Fig. 1). In prism-reared owls that have acquired a shifted map of ITD in the ICX, this topographic projection is altered in an adaptive manner (Fig. 4).

Anatomical remodeling has been demonstrated with both retrograde and anterograde tracers. Retrograde tracers injected into the ICX label large numbers of neuronal somata not only in the normal location in the ICC, but also in an abnormal location in the ICC where, normally, few somata are labeled (30). ITD tuning at this location in the ICC matches the adaptively shifted ITD tuning expressed at the site of injection in the ICX. Consistent with this result, anterograde tracers injected into the ICC label numerous, branching, bouton-laden axons not only in the normal projection zone in the ICX, but also in a zone of the ICX that normally contains few labeled processes (Fig. 3A *Right*). The adaptively shifted ITD tuning of ICX neurons in this zone matches the tuning at the site of the ICC injection.

The experience-dependent change in the axonal projection from the ICC to the ICX could result either from differential sculpting of an initially broad and dense projection field or from

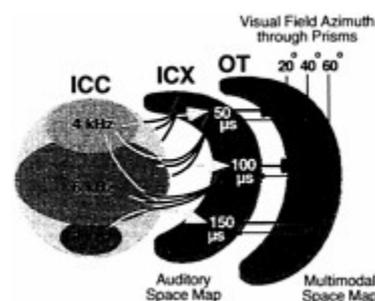


Fig. 4. Schematic representation of the change in the anatomical projection from the ICC to the ICX that results from early prism experience, as indicated by the data shown in Fig. 3. For each direction of prismatic displacement, an abnormal rostralward projection of ICC axons (red arrows) appears on one side of the brain and a caudalward projection appears on the other (data not shown). In addition, the normal projection persists (gray arrows). For additional details, see Fig. 1.

axonal elaboration in the adaptive portion of the projection field. Injections of anterograde tracer into the ICC of normal juveniles (at the age when prism experience begins) reveal that the projection in juveniles is indeed broader than in adults (W.M.D., D.Feldman, and E.I.K., unpublished data). However, the distribution of axonal densities in the juvenile projection field is far from sufficient to account for the axonal density in the adaptive portion of the projection field in prism-reared adults (Fig. 3B). Therefore, new axons must be formed. Moreover, because all axons in the adaptive portion of the projection field are studded with synaptic boutons, prism experience must result in synapto-genesis as well as in axonal elaboration.

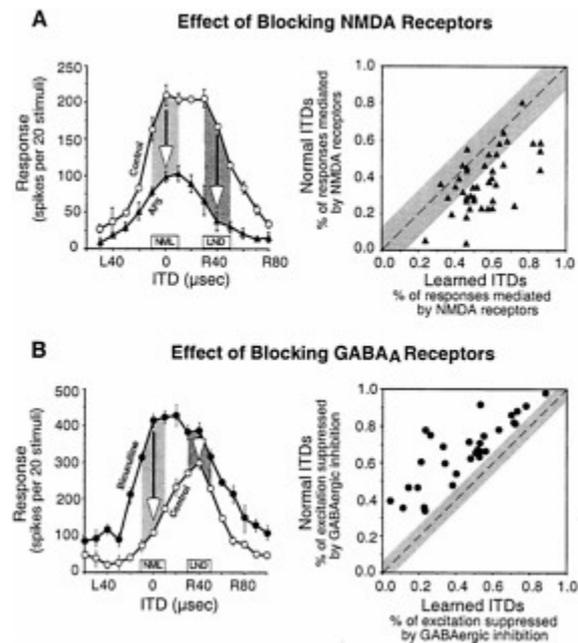


Fig. 5. Pharmacological traces of learning in the ICX of the barn owl. (A) Newly learned responses are differentially mediated by NMDA receptors. (Left) Effect of iontophoretic application of AP5, a specific antagonist for NMDA receptors, on the ITD tuning of an ICX site that was at an early stage of learning. Arrows indicate the AP5-induced decrease in responses to the normal (NML) and learned (LND) ITDs, respectively; learned ITDs were 40 μ s away toward right-ear leading from normal ITDs. (Right) Comparison of the proportion of the responses mediated by NMDA receptors, calculated as (control response—response with AP5)/control response, between normal and learned ITDs for all ICX sites ($n=41$) that were at an early stage of learning. The diagonal dashed line denotes equal percentage of NMDA-receptor-mediated responses to normal and learned ITDs. The gray region represents the mean ± 2 SD of the difference in the percentage of NMDA-receptor-mediated responses to ITDs ± 40 μ s away from the best ITD on each side of ITD tuning curves in normal owls. Data are from Feldman and Knudsen (34). (B) Normal responses are selectively suppressed by GABAergic inhibition. (Left) Effect of iontophoretic application of bicuculline, a specific antagonist for GABA_A receptors, on the ITD tuning of an ICX site that was at a late stage of learning. Arrows indicate the amount of excitation that was suppressed by GABAergic inhibition at the normal (NML) and learned (LND) ITDs, respectively. (Right) Comparison of the percentage of excitation that was suppressed by GABAergic inhibition, calculated as (response with bicuculline—control response) / response with bicuculline, between normal and learned ITDs for all ICX sites ($n=30$) that were at a late stage of learning. The diagonal dashed line denotes an equal percentage of GABAergic suppression for the normal and learned ITDs. The gray region represents the mean ± 2 SD of the difference in percentage of GABAergic suppression of excitation for ITDs ± 40 μ s away from the best ITD on each side of ITD tuning curves in normal owls. Data are from Zheng and Knudsen (35).

The acquisition of the learned, adaptive circuitry does not seem to come at the expense of the normal circuitry: The density of axons in the normal portion of the projection field is about the same in prism-reared and normal owls (Fig. 3B). Thus, at the anatomical level, the normal circuitry is able to coexist with the newly acquired circuitry.

Pharmacological Traces of Learning

Pharmacological experiments, in which specific neurotransmitter receptors are blocked in the ICX of prism-reared owls, confirm that the ICX is a site of plasticity and reveal the roles of certain classes of receptors in the adaptive, functional changes that take place.

The *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptors is particularly important for the expression of newly learned responses in the ICX. In the ICX of normal owls, blocking NMDA receptors, with iontophoretic injections of D,L-2-amino-5-phosphonopentanoic acid (AP5), reduces auditory responses by more than 50%, on average, across the entire range of effective ITDs (32). In the ICX of prism-reared owls in which learned responses are becoming strong, blocking NMDA receptors has a differentially greater effect on the newly learned responses and sometimes completely eliminates them (Fig. 5A) (33, 34). In contrast, nonspecific glutamate receptor antagonists do not have this differential blocking effect on the expression of learned responses.

The A subtype of γ -aminobutyric acid (GABA) receptors plays an essential role in suppressing inappropriate normal responses in an ICX that is expressing a fully shifted map of space (35). To enable a full, adaptive shift of an ICX tuning curve,

responses to normal ITDs must be eliminated. Iontophoretic injection of the GABA_A receptor blocker bicuculline at an ICX site that is expressing a fully shifted ITD tuning curve results in an immediate reexpression of strong normal responses and a shift of the tuning curve back toward normal values of ITD (Fig. 5B). Thus, consistent with the preservation of normal anatomical circuitry in prism-reared owls discussed above, functional connections mediating normal responses remain strong (although slightly weaker than normal) in an ICX that is expressing a shifted map of ITD, but responses to those inputs are differentially suppressed by GABAergic inhibition.

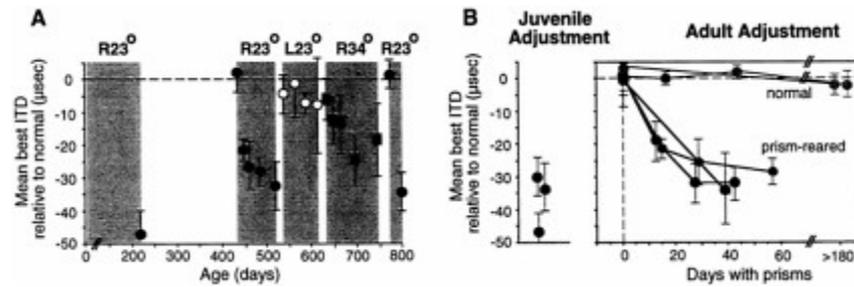


Fig. 6. Endurance of the traces of learning in the midbrain auditory localization pathway of the barn owl. Juvenile owls adjusted to R23° prisms. The prisms were then removed for a period of at least 6 months. The ability of the ITD map in the optic tectum to shift in response to prisms was tested again after the owls reached adulthood. Barn owls reach adulthood at about 7 months old. (A) Adjustments of tectal ITD tuning (measured as shown in Fig. 2C) in an adult owl that had adjusted to R23° prisms as a juvenile. The plot shows mean and standard deviation of the shift in best ITDs relative to predicted normal, measured at various ages. Shaded regions represent periods of experience with prisms of different strength or direction, as indicated above. Adaptive adjustment to L23° prisms would have been above the 0- μ s line. Regardless of the strength or direction of the prisms, adult adjustment is limited to the range of juvenile adjustment. (B) Shift of ITD tuning in three prism-reared (black) versus two normal (gray) adults. The two normal adults had no prior experience with prisms. The three prism-reared adults had adjusted to R23° prisms previously as juveniles (Left). All adult owls were over 1 year old. Data are from Knudsen (37).

Traces of Juvenile Learning Endure into Adulthood

Large changes in auditory spatial tuning are induced readily by manipulation of experience in young animals (36). In adult animals, adaptive adjustments may occur, but the range of adjustment is restricted. The period during development when large-scale changes occur in response to prism experience is referred to as the sensitive period for this manipulation (36). Within the sensitive period, horizontally displacing prisms can cause shifts of up to 70 μ s in the ITD tuning of tectal neurons. For owls past the sensitive period, the magnitude of adjustment for first-time exposure to identical prisms is substantially lower: Mean shifts in ITD tuning rarely exceed 10 μ s in naïve adults (Fig. 6B, normal).

In contrast, prism-reared adults exhibit an enhanced capacity for adaptive adjustments. Restoration of normal experience (by prism removal) leads to the recovery of normal ITD tuning in the optic tectum, and this capacity to recover normal tuning persists throughout life (Fig. 6A and ref. 36). In addition, reexposure to prisms leads to the reexpression of the adaptive responses that were learned during the sensitive period (37). For example, in an adult in which mean ITD tuning had shifted by more than 30 μ s toward left-ear leading values as a juvenile, and had subsequently shifted back to near 0 μ s after a period of normal experience, prism reexposure leads to a shift in mean ITD tuning of 30 μ s toward left-ear leading values (Fig. 6A). This enhanced capacity of plasticity persists well into adulthood. However, adult experience with a larger prismatic displacement does not lead to tuning shifts to larger left-ear leading values, and adult experience with prisms that displace the visual field in the opposite direction does not lead to tuning shifts to right-ear leading values (Fig. 6A). Thus, the acquisition of alternative tuning during the sensitive period leaves an enduring trace in this pathway that enables substantial functional plasticity in adults. This trace of learning specifically reflects the range of tuning states acquired during the juvenile period.

The persistence of a trace of learning in the midbrain localization pathway is characteristic of sensitive period learning, as evidenced by similar effects of song learning in song birds (38), imprinting in birds and mammals (39), and language learning in humans (40).

Comparison with Plasticity in Other Systems

Functional and anatomical plasticity, analogous to that described here for the midbrain localization pathway of the barn owl, has been observed in many other pathways and in many other species. In the hippocampus (41), visual and somatosensory cortices (42–46), and optic tectum of other species (47), for example, NMDA receptor-mediated excitation is critically involved in the acquisition of new responses, GABAergic inhibition regulates the plasticity, and patterns of axonal projection can be modified.

In most of these systems, the degree to which the plastic changes are adaptive is unclear. Abnormal experience often takes the form of deprivation, denervation, or excessive use (42, 46). Although the plastic changes that result from such manipulations are likely to be adaptive, the degree to which this is true is difficult to ascertain, because the importance of the neuronal changes for behavior is unknown.

In contrast, the adaptive nature of the plasticity in the midbrain localization pathway is clear. The role of the midbrain pathway in mediating orienting behavior is established (48). Moreover, orienting behavior is readily quantified, and the tuning of neurons for specific auditory cues is both precise and predictable (e.g., Fig. 2). Together, these properties allow quantitative links to be formed between behavioral learning and neuronal plasticity. These links reveal the adaptive nature of the functional and anatomical changes in this system. In addition, unlike plasticity reduced by deprivation, denervation, or excessive use (which may result simply from the induced imbalances in afferent activity and may be explained on the basis of self-organizational principles), the plasticity described here for the auditory localization pathway is instructed (49). Analogous to instructed motor learning in the cerebellum (50), an instructive signal (visually based, in the case of prism experience) controls the adaptive adjustment of the auditory space map in the midbrain.

One unusual property of the plasticity in the midbrain localization pathway is the low level of competition between alternative afferent pathways to the ICX. In the visual and somatosensory cortices of mammals, afferent inputs compete anatomically and physiologically for representation (43, 44, 46). In contrast, axonal projections corresponding to normal and learned inputs coexist in an ICX expressing a fully shifted map of space (Fig. 3B), and excitatory drive conveyed through both of these inputs is strong (35). The low level of competition between alternative afferent pathways may be critical for allowing the long-term retention of alternative maps despite disuse, enabling them to be restored at a later time should they become adaptive (Fig. 6).

A close parallel to the functional plasticity described here for the midbrain localization pathway in barn owls is found in the analogous pathway in mammals. In mammals, the homologue of the optic tectum, the superior colliculus, guides orienting movements of the eyes and head (9, 48). As in owls, it contains a multimodal map of space, and the auditory space map is shaped by experience during early life (18, 31). In ferrets, the auditory space map is altered adaptively in response to abnormal auditory experience or surgical rotation of the eye (analogous in most respects to prism experience). Although the sites and mechanisms of plasticity in this pathway in mammals are as yet unknown, it is likely that they share many features in common with those described here for the barn owl.

The properties of the midbrain localization pathway in the barn owl allow questions about the characteristics and mechanisms of learning to be addressed at many levels, from behavioral to molecular. Because of the tight links in this system among behavior, neuronal function, and neuronal structure, behavioral results can directly inform physiological and anatomical experiments, and vice versa. Consequently, a multilevel approach in this system greatly facilitates the search for principles and mechanisms of learning.

We are grateful for the substantial contributions of Drs. Michael Brainard and Dan Feldman to the work described here. We thank Phyllis Knudsen for technical assistance during many of these projects and for help in preparing the figures, and P. Hyde, B. Linkenhoker, and Dr. Y. Gutfreund for helpful comments on the manuscript. This work is supported by the National Institute of Deafness and Other Communication Disorders, National Institutes of Health Grant RO1 DC00155-18 and by a McKnight Senior Investigator Award.

1. Wightman, F.L. & Kistler, D.J. (1992) *J. Acoust. Soc. Am.* **91**, 1648–1661.
2. Knudsen, E.I. & Konishi, M. (1979) *J. Comp. Physiol.* **133**, 13–21.
3. Middlebrooks, J.C., Makous, J.C. & Green, D.M. (1989) *J. Acoust. Soc. Am.* **86**, 89–108.
4. Keller, C.H., Hartung, K. & Takahashi, T.T. (1998) *Hear. Res.* **118**, 13–34.
5. Middlebrooks, J.C. (1992) *J. Acoust. Soc. Am.* **92**, 2607–2624.
6. Cohen, Y.E. & Knudsen, E.I. (1999) *Trends Neurosci.* **22**, 128–135.
7. Carr, C.E. & Konishi, M. (1990) *J. Neurosci.* **10**, 3227–3246.
8. Hyde, P.S. & Knudsen, E.I. (2000) *J. Comp. Neurol.* **421**, 146–160.
9. Freedman, E.G., Stanford, T.R. & Sparks, D.L. (1996) *J. Neurophysiol.* **76**, 927–952.
10. Knudsen, E.I. & Knudsen, P.F. (1996) *Exp. Brain Res.* **108**, 23–32.
11. Rauschecker, J.P. (1998) *Audiol. Neurootol.* **3**, 86–103.
12. Knudsen, E.I. & Knudsen, P.F. (1996) *Nature (London)* **383**, 428–431.
13. Knudsen, E.I., Esterly, S.D. & du Lac, S. (1991) *J. Neurosci.* **11**, 1727–1747.
14. Hofman, P.M., Van Riswick, Jos, G.A. & Van Opstal, A.J. (1998) *Nat. Neurosci.* **1**, 417–421.
15. Miller, G.L. & Knudsen, E.I. (1999) *J. Neurosci.* **19**, 2326–2336.
16. Gold, J.I. & Knudsen, E.I. (2000) *J. Neurosci.* **20**, 3469–3486.
17. Knudsen, E.I. (1999) *J. Comp. Physiol.* **185**, 305–321.
18. King, A.J. (1993) *Exp. Physiol.* **78**, 559–590.
19. Knudsen, E.I., Knudsen, P.F. & Esterly, S.D. (1982) *Nature (London)* **295**, 238–240.
20. Knudsen, E.I. (1983) *Science* **222**, 939–942.
21. Mogdans, J. & Knudsen, E.I. (1992) *J. Neurosci.* **12**, 3473–3484.
22. Knudsen, E.I., Esterly, S.D. & Olsen, J.F. (1994) *J. Neurophysiol.* **71**, 79–94.
23. Gold, J.I. & Knudsen, E.I. (1999) *J. Neurophysiol.* **82**, 2197–2209.
24. Gold, J.I. & Knudsen, E.I. (2000) *J. Neurosci.* **20**, 862–877.
25. Knudsen, E.I. & Brainard, M.S. (1991) *Science* **253**, 85–87.
26. Brainard, M.S. & Knudsen, E.I. (1995) *J. Neurophysiol.* **73**, 595–614.
27. Takahashi, T. & Konishi, M. (1986) *J. Neurosci.* **6**, 3413–3422.
28. Brainard, M.S., Knudsen, E.I. & Esterly, S.D. (1992) *J. Acoust. Soc. Am.* **91**, 1015–1027.
29. Brainard, M.S. & Knudsen, E.I. (1993) *J. Neurosci.* **13**, 4589–4608.
30. Feldman, D.E. & Knudsen, E.I. (1997) *J. Neurosci.* **17**, 6820–6837.
31. King, A.J., Hutchings, M.E., Moore, D.R. & Blakemore, C. (1988) *Nature (London)* **332**, 73–76.
32. Feldman, D.E. & Knudsen, E.I. (1994) *J. Neurosci.* **14**, 5939–5958.
33. Feldman, D.E., Brainard, M.S. & Knudsen, E.I. (1996) *Science* **271**, 525–528.
34. Feldman, D.E. & Knudsen, E.I. (1998) *J. Neurosci.* **18**, 3073–3087.
35. Zheng, W. & Knudsen, E.I. (1999) *Science* **284**, 962–965.
36. Brainard, M. & Knudsen, E. (1998) *J. Neurosci.* **18**, 3929–3942.
37. Knudsen, E.I. (1998) *Science* **279**, 1531–1533.
38. Doupe, A.J. & Kuhl, P.K. (1999) *Annu. Rev. Neurosci.* **22**, 567–631.
39. Hess, E.H. (1973) *Imprinting: Early Experience and the Developmental Psychobiology of Attachment* (Van Nostrand Reinhold, New York).
40. Tong, Y.C., Busby, P.A. & Clark, G.M. (1988) *J. Acoust. Soc. Am.* **84**, 951–962.
41. Martin, S.J., Grimwood, P.D. & Morris, R.G.M. (2000) *Annu. Rev. Neurosci.* **23**, 649–711.
42. Buonomano, D.V. & Merzenich, M.M. (1998) *Annu. Rev. Neurosci.* **21**, 149–186.
43. Daw, N.W. (1994) *Invest. Ophthalmol. Visual Sci.* **35**, 4168–4179.
44. Katz, L.C. & Shatz, C.J. (1996) *Science* **274**, 1133–1138.
45. Feldman, D.E. (2000) *Nature Neurosci.* **3**, 303–304.
46. Kaas, J.H. (2000) *The New Cognitive Neuroscience* (MIT Press, Cambridge, MA).
47. Udin, S.B. & Grant, S. (1999) *Prog. Neurobiol.* **59**, 81–106.
48. Stein, B. & Meredith, M. (1993) *The Merging of the Senses* (MIT Press, Cambridge, MA).
49. Knudsen, E.I. (1994) *J. Neurosci.* **14**, 3985–3997.
50. Raymond, J.L., Lisberger, S.G. & Mauk, M.D. (1996) *Science* **272**, 1126–1131.

PLASTICITY IN THE NEURAL CODING OF AUDITORY SPACE IN THE MAMMALIAN BRAIN

Andrew J.King*, Carl H.Parsons, and David R.Moore

University Laboratory of Physiology, Oxford University, Parks Road, Oxford OX1 3PT, United Kingdom

Sound localization relies on the neural processing of monaural and binaural spatial cues that arise from the way sounds interact with the head and external ears. Neurophysiological studies of animals raised with abnormal sensory inputs show that the map of auditory space in the superior colliculus is shaped during development by both auditory and visual experience. An example of this plasticity is provided by monaural occlusion during infancy, which leads to compensatory changes in auditory spatial tuning that tend to preserve the alignment between the neural representations of visual and auditory space. Adaptive changes also take place in sound localization behavior, as demonstrated by the fact that ferrets raised and tested with one ear plugged learn to localize as accurately as control animals. In both cases, these adjustments may involve greater use of monaural spectral cues provided by the other ear. Although plasticity in the auditory space map seems to be restricted to development, adult ferrets show some recovery of sound localization behavior after long-term monaural occlusion. The capacity for behavioral adaptation is, however, task dependent, because auditory spatial acuity and binaural unmasking (a measure of the spatial contribution to the “cocktail party effect”) are permanently impaired by chronically plugging one ear, both in infancy but especially in adulthood. Experience-induced plasticity allows the neural circuitry underlying sound localization to be customized to individual characteristics, such as the size and shape of the head and ears, and to compensate for natural conductive hearing losses, including those associated with middle ear disease in infancy.

Considerable plasticity exists in the neural circuits that process sensory information. Although plasticity is greatest during development, certain aspects of the mature brain maintain the capacity to reorganize in response to changes in the activity patterns of sensory inputs. Experience-mediated plasticity is most commonly associated with higher-level response properties that are generated within the brain by serial stages of computation. One of the best examples is provided by the neural mechanisms underlying sound localization, which are adapted by experience to the features of the individual.

Identifying the location of sounds produced by potential mates, prey, or predators is one of the most important functions of the auditory system. This ability relies on the extraction of direction-dependent cues generated by the head, torso, and external ears. For localization in the horizontal plane, the separation of the two ears allows mammals to use interaural time differences (ITDs) at low frequencies and interaural level differences (ILDs) at high frequencies (1, 2). These binaural cues also contribute to the ability of listeners to detect and discriminate signals of interest against a background of masking noise (3).

ITDs and ILDs do not, by themselves, provide a sufficient basis for localizing a sound source. For individual frequencies, both cues are spatially ambiguous in that the same binaural disparity value may arise from different directions in space. Moreover, when binaural cues are introduced into sounds presented over headphones, humans report that the sound is lateralized inside the head, rather than localized to a specific direction in the external environment. Additional spatial cues are generated, however, when broadband sounds travel from a source in real space and pass through the external ear. Depending on the direction of the sound source, certain frequency components will be increased or decreased in amplitude, giving rise to specific spectral patterns at each ear (1, 2). Along with the acoustic shadow cast by the head, these spectral cues contribute to the complex pattern of ILDs that are generated in the free field and are responsible for resolving front-back confusions and for localization in the vertical plane, where minimal binaural information is available. Presentation over headphones of virtual acoustic space signals, which incorporate all of the spatial information associated with free-field sound sources, including the filtering characteristics of the external ears, have highlighted the role of spectral cues in producing an externalized perception of auditory space (4). Monaural spectral cues are also presumably responsible for the capacity of unilaterally deaf humans to localize sounds in both azimuth and elevation (5, 6)

Neural Representations of Auditory Space

The initial processing of ITDs and ILDs is carried out predominantly, but not exclusively, in the medial superior olive and the lateral superior olive, respectively (7). These nuclei project to the central nucleus of the inferior colliculus (IC) in the midbrain, which also receives inputs from the contralateral dorsal cochlear nucleus, where monaural spectral cues seem to be processed (8), and from other auditory brainstem nuclei. Although the multiple inputs to the central nucleus of the IC allow further processing to take place (9–11), neurons sensitive to different auditory localization cues remain largely segregated within this nucleus, which reflects the frequency dependence of the cues and the tonotopic organization of the brainstem auditory nuclei (7, 12). Spatial information is then relayed via the thalamus to the primary auditory cortex. The importance of the primary auditory cortex in sound localization is indicated by the inability of mammals to approach sound sources in the contralateral hemifield after unilateral lesions of this cortical area (13, 14). How sound source location is encoded in the primary auditory cortex remains unclear but is likely to be based on the distribution of activity, in terms of both firing rate and temporal discharge pattern, across assemblies of neurons (15).

In addition to more complex behaviors that involve the cortex, an important function of auditory localization is the guidance of reflexive orienting responses that shift attention toward unexpected sounds in the environment. These seem to be mediated

*To whom reprint requests should be addressed. E-mail: ajk@physiol.ox.ac.uk.

This paper was presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: IC, inferior colliculus; ILD, interaural level difference; ITD, interaural time difference; MAA, minimum audible angle; MED, middle ear disease; SC, superior colliculus.

at the level of the superior colliculus (SC), which receives inputs from the external nucleus of the IC, from the nucleus of the brachium of the IC, and from other auditory regions of the brainstem and cortex (16, 17). The SC is of particular interest to studies of sound localization, because, in contrast to the primary auditory cortex, the spatial selectivity of the auditory neurons found there varies systematically to form a map of space (18–20), much like the visual and tactile maps that also exist in the SC and elsewhere in the brain.

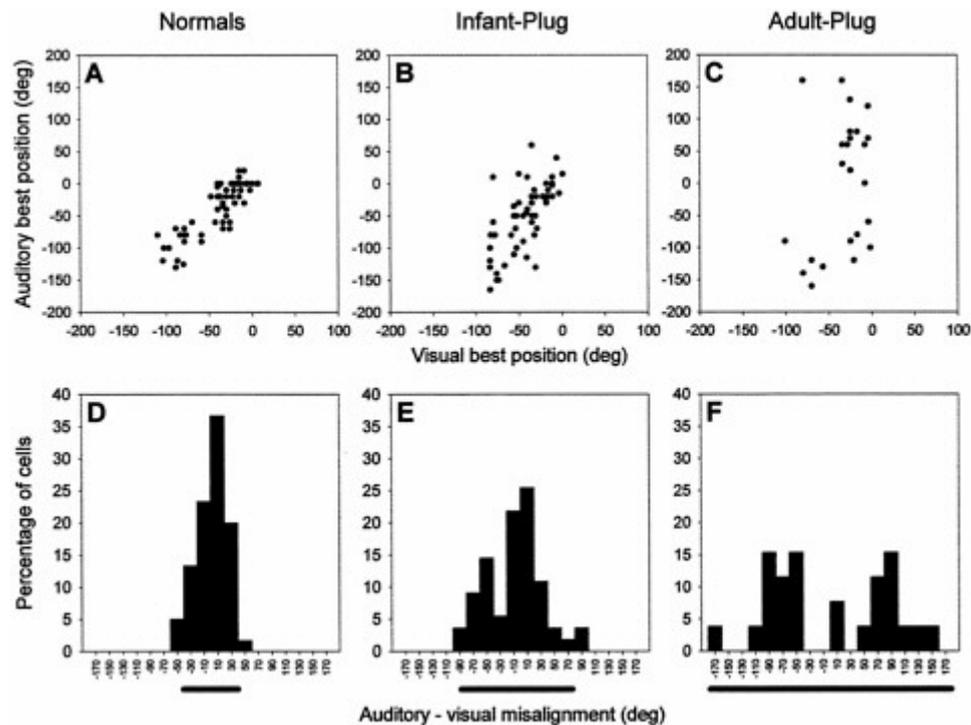


Fig. 1. Effects of chronic monaural occlusion on the registration of the auditory and visual maps in the ferret SC. Each panel shows the relationship between the representations of visual azimuth in the superficial layers and auditory azimuth in the deeper layers of the SC. (A–C) Recordings were made from anesthetized ferrets, and for each vertical electrode penetration, the visual best azimuth is plotted against the auditory best azimuth (measured with 100-ms broadband noise bursts at sound levels of 25–35 dB above unit threshold). deg, degree. (D–F) The frequency histograms plot the angular difference between the visual and auditory best azimuths; the bar below each histogram is centered on the mean misalignment and extends to 2 SDs on either side. (A and D) Data from normal, adult ferrets. (B and E) Data from adult ferrets that had been raised from just before the onset of hearing (which in ferrets occurs \approx 4 weeks after birth) with the ear ipsilateral to the recording site occluded. (C and F) Data from adult ferrets that had one ear plugged for a comparable period, this time beginning when they were at least 6 months old. The data shown in B, C, E, and F were obtained with the earplug still in place. F tests revealed that the variance in auditory-visual misalignment is significantly different between each of the three groups. Because the superficial layer visual map showed a high degree of topographic order in each case, the increased scatter in the relationship between the two maps in the plugged animals is indicative of poorer topographic order in the auditory representation, as adults. Data from juvenile animals are based on those in ref. 36.

Whereas neural maps of visual space and of the body surface arise from topographically ordered projections from the primary afferents innervating the retina and skin, respectively, the map of auditory space has to be computed by tuning neurons to combinations of localization cues that correspond to specific sound directions. In mammals, a systematic variation in sensitivity to ILDs and to monaural spectral cues seems to provide the acoustical basis for the map of auditory space in the SC (1). Given the high-frequency bias of SC neurons, it is less likely that ITDs are involved. Nevertheless, ITD sensitivity has been demonstrated for cat SC neurons over ranges exceeding those produced naturally by the separation of the ears (21). Because the response latencies of auditory neurons tend to be shorter at higher sound levels, the timing of inputs arriving from each ear may differ for high-frequency sounds located to one side of the head. Consequently, apparent sensitivity to ITDs may provide a basis for the encoding of ILDs.

Auditory, visual, and somatosensory inputs converge on the deeper layers of the mammalian SC, whereas the superficial layers are exclusively visual in function. Despite this laminar segregation, the sensory inputs to both regions are topographically aligned, as illustrated in Fig. 1 A and D for the visual representation in the superficial layers and the auditory representation in the deeper layers. Individual neurons in the deeper layers of the SC often receive inputs from more than one sensory

modality, and their responses are determined by the spatiotemporal relationship between different sensory stimuli (22). The responses of these neurons tend to be enhanced when different sensory cues are presented in close spatial and temporal proximity, as would be the case for a source that can be both seen and heard, but depressed with stimuli that are widely separated in space or time. These multisensory interactions highlight the importance of aligning the different sensory maps in the SC. Indeed, modulation of auditory responses by eye-position signals may allow intersensory map registration to be largely preserved, even after gaze shifts in awake mammals that would otherwise misalign the visual and auditory maps (23–25). The correspondence between sensory and motor maps in the SC also provides an efficient way of using spatial information from different modalities to guide orienting movements toward novel stimuli.

Learning to Localize with Your Own Ears

Coordinating the neural maps of space for different sensory modalities involves matching specific values of the auditory localization cues to positions on the retina and the body surface. However, the monaural and binaural cue values corresponding to particular directions in space depend on the size, shape, and separation of the ears and may vary markedly between different individuals (26) and even between the left and right sides of the same subject (27). The cue values also undergo substantial changes during development as the head and external ears grow (1, 28, 29). These changes may contribute to the steady improvement in the ability of human infants to localize sound with increasing age (30). Similarly, the map of auditory space in the SC matures gradually during postnatal life, with normal topographic order emerging at about the same stage that the monaural spectral cues and ILDs attain their adult values (1). It therefore seems unlikely that the neural mechanisms responsible for processing spatial information are adjusted continuously during development to adapt to the changing cues. Nevertheless, adult human listeners localize virtual acoustic space stimuli derived from the head-related transfer functions of their own ears more accurately than those based on the ears of other subjects (31, 32). This finding suggests that neural representations of auditory space are calibrated by experience of the spatial cues provided by the subject's own ears.

Experience Shapes the Development of the Auditory Space Map in the SC

The map of auditory space in the SC has provided a very useful system for investigating experience-dependent plasticity in the neural coding of sound source location. By manipulating the sensory cues available during development, studies in barn owls (33) and mammals (34) have demonstrated the importance of both auditory and visual experience in shaping the spatial tuning of SC neurons.

Experiments in which the visual field was displaced with respect to the head, either optically in owls by prism rearing (35) or surgically in young ferrets by removal of one of the extraocular muscles (34, 36), revealed compensatory shifts in the auditory space map, suggesting that visual cues play a dominant role in establishing intersensory map registration. Much progress has since been made in owls in identifying the basis for the visual calibration of auditory spatial tuning and the time scale over which prism-induced changes can be induced (33). The site and mechanisms of plasticity have received less attention in mammals, although lesion studies suggest that the guiding visual signals may arise during development from the superficial layers of the SC (37). These exclusively visual layers project topographically both to the deeper layers of the SC and to the nucleus of the brachium of the IC, which itself contains a coarse map of sound azimuth (38).

Consistent with the guiding role of vision is the finding that degradation of visual cues during development leads to some abnormalities in the auditory representation. For example, raising ferrets with their eyelids sutured results in a higher incidence of auditory units that are ambiguously tuned to two different sound directions, although otherwise the map seems to be essentially normal (39). The ability of visually deprived ferrets to judge the location of sound sources in the horizontal plane is also no different from that of normally sighted ferrets. In keeping with the SC results, however, these animals do exhibit more front-back confusions than normal controls. On the other hand, we found that visually deprived ferrets actually show an improvement in auditory spatial acuity at lateral sound locations (40), a result that has also been reported in localization studies of lid-sutured cats (41) and blind humans (42). The conclusion from these studies is that, within a multisensory environment, the more reliable and accurate cues available through the visual system are used to coordinate spatial information provided by the auditory system. This ability to combine information across the senses leads to enhanced stimulus detection, localization, and identification as well as faster reactions after stimulus presentation (22). However, in the absence of vision, compensatory changes can take place within the brain in the processing of auditory and other sensory signals.

Auditory map plasticity has also been examined by using experimental procedures that alter the acoustic localization cues available (33, 34). Studies in mammals suggest that if the spatial cues are sufficiently degraded during development, then the map of auditory space in the SC will fail to form properly. For example, removal of the pinna and concha of the external ear disrupts the spectral localization cues such that, compared with those of normal animals, they provide much less information for discriminating between anterior and posterior sound directions (43, 44). Bilateral removal of the pinna and concha in juvenile ferrets impairs both the emergence of topographic order in the auditory space map (43) and the ability of mature animals to localize broadband sounds (44).

An adaptive change in the auditory space map has been observed, however, by raising owls (45) and ferrets (36) with unnatural binaural cues produced by plugging one of the ears. Fig. 1*B* and *E* shows the relationship between the auditory map in the deeper layers and the visual map in the superficial layers of the SC in adult ferrets that had been raised with one ear occluded. These data, which were obtained with the earplug in place, show that the correspondence between the two maps is not as good as in normal ferrets (Fig. 1*A* and *D*) but significantly better than in another group of animals that received a similar period of monaural occlusion as adults (Fig. 1*C* and *F*).

In owls, the basis for this adaptation involves a compensatory shift during development in the tuning of auditory neurons toward the abnormal ITDs and ILDs produced by the earplug (46). Consequently, removal of the plug results in an immediate misalignment of the auditory and visual receptive fields, which, in juvenile owls, gradually decreases over a period of several weeks (45).

Irrespective of whether the ferrets received an earplug in infancy or as adults, we observed closer visual-auditory map alignment after plug removal. This result suggests that monaural occlusion in adult ferrets does not alter the sensitivity of SC neurons to the localization cues, as normal topographic order in the representation is apparent only with a balanced binaural input. The basis for the adaptation observed in the auditory space map in the infant-plugged ferrets is presently unclear, but, as discussed below, the neural plasticity involved is also likely to underlie the effects of early monaural occlusion on sound localization behavior.

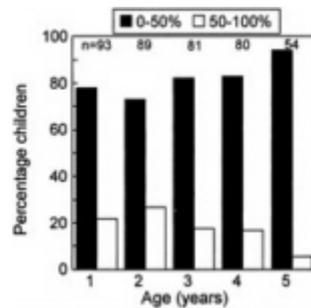


Fig. 2. Incidence of MED in the first 5 years of life (from a study performed in Oxford, 1993–1998). A group of children with normal birth histories were examined every month from birth with bilateral otoscopy and tympanometry to detect the occurrence of fluid in the middle ear (see ref. 47 for further details). The histogram bars show, for each year, the percentage of children who had fluid, in one or both ears, on less than (black bars) or more than (white bars) 50% of the examinations in that year. Most of the children who had high rates of MED in 1 year also had high rates in the other years. Overall, about 15% of the children in our sample had MED for more than half of the first 5 years of their lives (S.C.Hogan and D.R.M., unpublished data).

Behavioral Evidence for Plasticity of Sound Localization

In addition to allowing for growth-related changes in localization cues, plasticity of the auditory system allows for some degree of adaptation to the altered auditory input produced by a conductive hearing loss. A common form of conductive loss, particularly prevalent in early childhood (ref. 47; Fig. 2), is middle ear disease (MED; also called otitis media with effusion). In MED, pressure changes associated with Eustachian tube malfunction result in the secretion of fluid into the middle ear cavity. This fluid impairs the impedance-matching function of the middle ear and both attenuates and delays sound passing to the cochlea (D.E.H. Hartley and D.R.M., unpublished observation). Because MED is usually either unilateral or bilaterally asymmetric, it changes the balance of input between the two ears, thereby generating abnormal binaural cues during infancy. There is therefore considerable interest in determining whether conductive hearing losses lead to changes in central auditory function, as suggested by the neurophysiological studies of the developing space map in the SC.

Research into the effects of MED on human hearing has been hampered by the usually poor data available on the precise history of the disease and, of course, by the lack of control over either the severity or the time course. To address these problems, we have studied the behavioral effects of monaural occlusion on various measures of localization in humans and ferrets. Earplugs produce sound attenuation and delays that are comparable to those produced by severe MED, but they do so in a far more controllable way.

Studies in adult humans have shown that, as expected, earplugs or other devices fitted to the external ear disrupt auditory localization, with differing degrees of adaptation reported during the variable and often relatively short periods used for manipulating the localization cues (6, 48–51). Studies in animals permit longer-term ear plugging, which can also be carried out during infancy, and have demonstrated considerable adaptation in certain circumstances. By measuring the accuracy of soundevoked head-orienting responses, Knudsen and colleagues (52) showed that juvenile owls gradually learn to reinterpret the abnormal spatial cues produced by plugging one of their ears. We examined the effects of monaural occlusion on the ability of ferrets to perform a spatial identification task, in which they were trained to approach the location of a sound source in the azimuthal plane (Fig. 3A). Ferrets that were raised with one ear occluded could perform this task just as accurately as normal controls, indicating that they had compensated for the presence of the earplug (Fig. 3B and C).

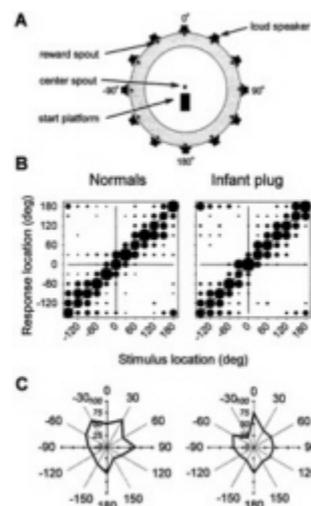


Fig. 3. Effects of chronic monaural occlusion in infancy on auditory localization. (A) Schematic view of chamber used to measure the spatial identification ability of ferrets. The animals were trained to stand on the start platform and initiate a trial by licking the center spout. Each trial consisted of a Gaussian noise burst (0–30 kHz, 100-ms duration) presented quasirandomly from 1 of 12 speakers placed at 30° intervals in the azimuthal plane. Within each testing session, five sound levels ranging randomly from 56 to 84 dB sound pressure level were used to minimize loudness cues. Ferrets were rewarded for approaching and licking the spout associated with the speaker that had been triggered. (B) Stimulus-response plots showing the combined data of three normal adult ferrets (Normals) and three ferrets that had been raised and tested with the left ear occluded with a plug that produced 30–50 dB attenuation (Infant plug). These plots illustrate the distribution of responses (ordinate) as a function of stimulus location (abscissa). The size of the dots indicates, for a given speaker angle, the proportion of responses made to different spout locations, deg, degree. (C) Polar plots showing the percentage of correct responses made to each of the 12 speaker locations. Comparison of the percentage scores and error magnitudes showed that ferrets reared with a plug in one ear perform as well as normally reared animals.

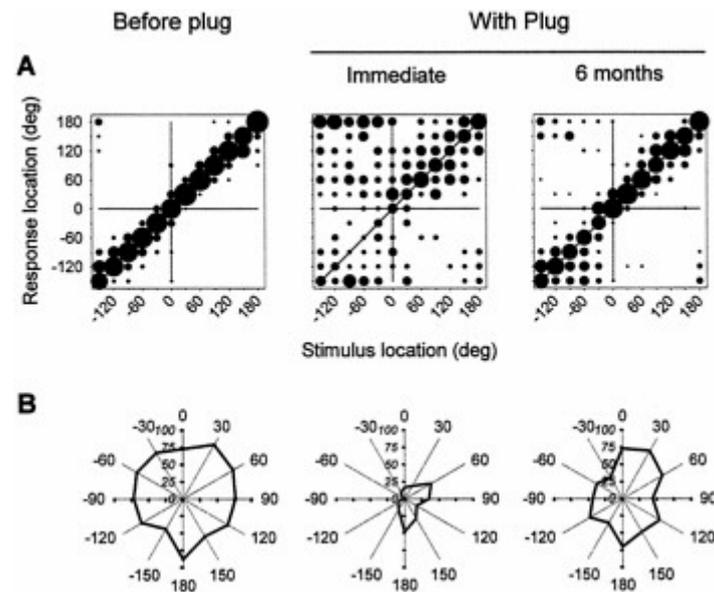


Fig. 4. Auditory localization data from three normally reared ferrets that had the left ear plugged in adulthood. These ferrets were tested extensively in the spatial identification task before ear plugging and therefore achieved higher scores than the normal control group illustrated in Fig. 3. Before plug panel (*Left*) shows performance of the animals with 100-ms noise bursts in the last test before insertion of the earplug. Data obtained when the earplug was in place (*With Plug*) are also shown (*Center and Right*). These tests were carried out during the day after insertion of the plug (*Immediate*) and again 6 months later (*6 months*). In the intervening period, the plug remained in place, but the ferrets were not tested. (*A*) Stimulus-response plots showing the distribution of errors at each speaker location, deg, degree. (*B*) Polar plots showing the percentage of correct responses at each speaker location. The ferrets initially performed poorly after occlusion of one ear in adulthood, even at much longer stimulus durations. Despite the improvement in performance observed overtime, these animals achieved significantly lower percentage scores and made larger localization errors with one ear plugged, compared with their performance before the ear was occluded.

Little after effect was apparent after plug removal, and the ferrets continued to localize almost just as accurately. Subsequently, the performance of these animals showed some improvement, and reinsertion of the earplug resulted in a small shift in their responses toward the side of the unplugged ear. Although consistent with an effect on binaural processing, these changes are much less marked than those produced by acute monaural occlusion in adult ferrets (Fig. 4). As with the effects of plugging one ear during infancy on the responses of ferret SC neurons, the adaptation to the earplug does not seem to be based simply on a systematic shift in the association between binaural cue values and directions in space. Instead, it may be that greater emphasis is placed on other cues that are not disrupted by ear plugging, a possibility that may also explain the adaptation reported in adult humans after monaural occlusion (48, 49).

Sound localization behavior in adult barn owls is not adjusted by long-term plugging of one ear (52). This result is in accordance with the earplug-induced plasticity of auditory spatial tuning in the midbrain, which, as shown in Fig. 1, seems to be regulated developmentally. However, our behavioral studies in ferrets have shown that auditory localization by adult ferrets does improve after several months of monaural occlusion, although not to the level exhibited by the same animals before plugging (Fig. 4). We have also observed a more rapid improvement in performance by training adult-plugged ferrets more extensively, but without feedback about the accuracy of performance, over the first few weeks after insertion of the earplug. These observations raise intriguing questions about the factors that contribute to plasticity of sound localization and the neural mechanisms by which they are brought about.

Effects of Conductive Hearing Loss on Other Measures of Spatial Hearing

Studies of natural conductive impairments in humans suggest that their effects on binaural hearing may depend on the nature of the task that is being measured (53). To determine what aspects of sound localization are altered by plugging one ear, we have examined the effects of monaural occlusion on two other measures of binaural hearing. We found that long-term monaural occlusion, either in infancy or in adulthood, impairs the ability of ferrets to discriminate between two broadband sounds in the azimuthal plane (Fig. 5). However, ferrets reared and tested with one ear occluded achieved significantly better scores than those plugged for a similar period as adults. This result suggests a degree of adaptation, at least during infancy, that may also explain the normal MAAs reported in humans with mild long-term conductive losses (5). The compensatory changes involved seem, however, to be much less complete than those observed in the spatial identification task.

We have also examined the effects of unilateral ear plugging in ferrets on free-field binaural unmasking (54). This measure of binaural hearing depends on interaural phase detection and underpins the ability of listeners to distinguish separate spatial streams of acoustic information. Binaural unmasking is probably the most common test used to assess binaural hearing in humans. Indeed, studies in children have shown that, after a history of

recurrent MED, binaural unmasking can remain below normal levels beyond the recovery of peripheral hearing (refs. 55 and 56; Fig. 6). In contrast to the effects of visual deprivation on the development of central visual function, however, the binaural consequences of long-term MED are reversible. Thus, children gradually reacquire normal hearing in the months to years after the cessation of MED (refs. 57 and 58; Fig. 6).

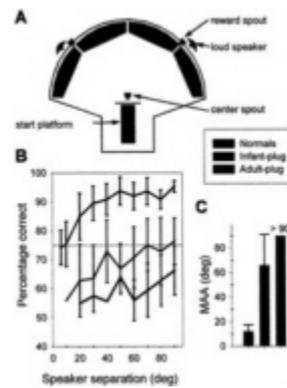


Fig. 5. Effects of chronic monaural occlusion on auditory spatial acuity. (A) Schematic view of chamber used to measure minimum audible angles (MAA) in ferrets. The animals were trained to initiate a trial by standing on the start platform and licking the center spout. This action triggered a noise burst (800 Hz to 24 kHz, 100-ms duration) from one of the two speakers placed symmetrically about the anterior midline. Ferrets were rewarded for licking the water spout in the same hemifield as the active speaker. The separation of the speakers was reduced between testing sessions according to the method detailed in ref. 44. Psychometric functions were measured by reducing the angular separation of the speakers from 90° to 6° or until performance did not differ from chance (56.5%, binomial distribution). (B) Mean (\pm SD) psychometric functions for normal ferrets (Normals, $n=6$), ferrets plugged in infancy (Infant-plug, $n=4$), and ferrets plugged as adults (Adult-plug, $n=3$). These data were obtained with the earplug in place, deg, degree. (C) MAAs (defined as the speaker separation corresponding to 75% correct performance) were calculated from the psychometric function data. Note that adult-plugged ferrets failed to reach 75% correct, even at the maximum speaker separation of 90°; their MAAs are therefore indicated by >90°.

Compared with normal adults, monaurally occluded ferrets achieved low levels of unmasking with the earplug in place, even after many months of plugging (Fig. 7). There was some indication that ferrets reared with one ear occluded showed less impairment than animals that had been plugged as adults (Fig. 7 B and C). However, in contrast to the localization experiments, these animals showed little sign of adaptation to the unilateral conductive hearing loss produced by the earplug. Like the children recovering from MED, they also had poor binaural unmasking for some months after removal of the earplugs.

These studies show that different aspects of spatial hearing are affected in different ways by a unilateral conductive hearing loss. Their findings are consistent with the notion that adaptation of sound localization by ferrets (and humans) may be caused less by compensatory changes in binaural hearing and more by a learned improvement in the use of other cues, such as monaural spectral cues, that are unaffected by the conductive hearing loss.

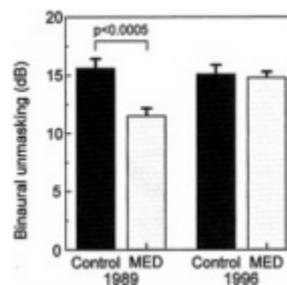


Fig. 6. Binaural hearing in children with a history of MED. Children aged 6–12 years, presenting at the Otolaryngology Department, Radcliffe Infirmary, Oxford, for tympanostomy surgery to relieve MED were paired, in 1989, with control children who had no known history of MED (from ref. 55). These two groups were examined for binaural unmasking—the difference in the detection threshold of a 500-Hz tone presented against a binaural noise stimulus when the tone is either in or out of phase between the two ears. This test is thought to provide a measure of the binaural contribution to the “cocktail party effect.” Children with a history of MED had impaired binaural hearing, even though their absolute thresholds were normal at the time of testing. These findings together suggest that a history of unilateral or asymmetric hearing loss can lead to impaired central auditory function. In 1996, 7 years later (from ref. 58), about two-thirds of the same children were reexamined. Both groups now had normal levels of binaural unmasking, suggesting that prolonged normal binaural experience can lead to the restoration of central auditory function.

Summary

Sound localization is calibrated by experience, particularly during development when head growth changes the relationship between auditory cues and directions in space. Studies of the maturation of intersensory map registration in the SC have shown that the auditory representation can accommodate abnormal auditory or visual cues induced by experimental manipulations of the sensory inputs. Behavioral correlates of these experience-induced changes in neuronal response properties are seen in the capacity of juvenile ferrets to adjust sound localization to the altered acoustic cues produced by chronic occlusion of one ear. The effects of plug removal suggest that, in both cases, the adaptation observed may be at least partly based on a greater dependence on monaural spectral cues. Other measures of binaural hearing show much less adaptation after long-term monaural occlusion. Indeed, given the similarity between the effects of plugging one ear in ferrets and MED in human infants on binaural unmasking, it would be of interest to examine whether children with a history of recurrent MED exhibit adaptive changes in their ability to localize sound.

Compensatory adjustments occurring at the level of the auditory space map in the SC in response to long-term monaural occlusion have thus far been observed only during early postnatal life, suggesting that the need for plasticity may decline once the head and ears have attained their adult size. Although behavioral adaptation to ear plugging is also most pronounced

during infancy, studies of sound localization in ferrets and humans indicate that the capacity of the brain to reinterpret auditory localization cues extends into adult life. Moreover, as with other recent examples of adult plasticity, training seems to facilitate this process.

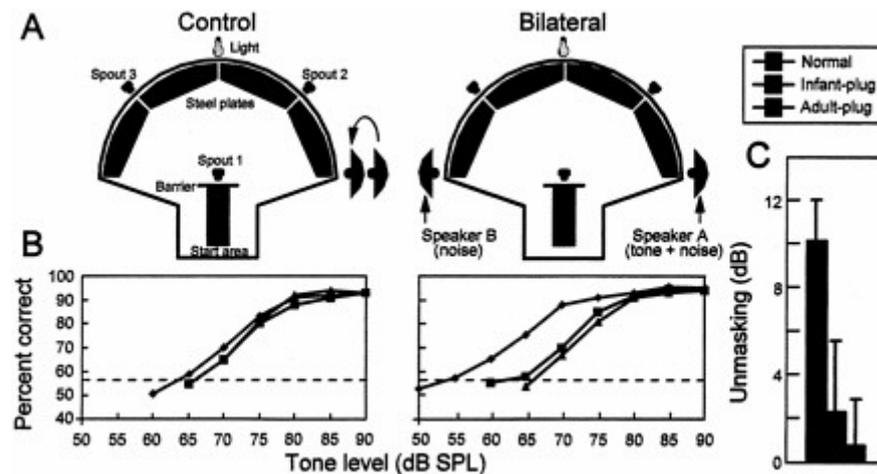


Fig. 7. Effect of chronic monaural occlusion on binaural unmasking in ferrets. (A) Adult ferrets were trained to detect a 500-Hz tone (delivered from speaker A) in the presence of a noise (delivered from speakers A and B) when a light flashed. The noise stimuli were presented continuously. On 50% of trials, a tone was presented when the ferret contacted spout 1. Success in the task was measured by the ferret correctly identifying the presence or absence of the tone (by going to spouts 2 or 3, respectively). In the control condition (tones and noise interaurally in phase), both speakers were positioned on the ferret's right side. In the bilateral condition, the noise was made out of phase with the tone by moving speaker B (noise alone) to the ferret's left side. The difference in tone threshold between the two conditions measured binaural unmasking. (B) Ferrets learned to perform this task at a high level; the dashed line shows performance that was statistically above chance (binomial distribution). However, normal ferrets, like normal humans (Fig. 6), consistently produced thresholds that were about 10 dB better in the bilateral condition, showing binaural unmasking. (C) After 3–12 months of unilateral (left) ear plugging, there was little or no difference in performance between the control and the bilateral conditions, while the plug was in place. This lack of difference was true both for ferrets plugged before the normal onset of hearing (Infant-plug) and for ferrets receiving equivalent experience in adulthood (Adult-plug). Note, however, that one of the four infant-plugged ferrets had 8 dB of residual unmasking. When the plug was finally removed, both groups had impaired unmasking, despite normal pure tone sensitivity in the previously plugged ear. Like humans who have had surgery for a conductive hearing loss (Fig. 6), the unplugged ferrets gradually recovered normal binaural unmasking over a period of several months (from ref. 54).

The auditory space map in the SC provides a very useful system for examining the neural basis for the plasticity of sound localization, at least in relation to orienting behavior. However, it is likely that attention will have to switch to the cortex if we are to understand fully the principles underlying adaptation of the more complex localization behaviors used in this study.

We are grateful to J.Schnupp for comments on this paper and to Jemma Hine, Sarah Hogan, Leslie Hulvershorn, Ze Dong Jiang, Oliver Kacelnik, Richard Lanyon, Hiroaki Matsuda, and Susan Spires for their contributions to the psychophysical studies. Our research is supported by the Wellcome Trust (through a Senior Research Fellowship to A.J.K.), by the Medical Research Council (D.R.M.), by Defeating Deafness, and by the Oxford McDonnell-Pew Centre for Cognitive Neuroscience.

- King, A.J. & Carlile, S. (1995) in *The Cognitive Neurosciences*, ed. Gazzaniga, M.S. (MIT Press, Cambridge, MA), pp. 279–293.
- Wightman, F.L. & Kistler, D.J. (1993) in *Human Psychophysics*, eds. Yost, W.A., Popper, A.N. & Fay, R.R. (Springer, New York), pp. 155–192.
- Moore, B.C.J. (1997) *An Introduction to the Psychology of Hearing* (Academic, San Diego).
- Wightman, F.L. & Kistler, D.J. (1989) *J. Acoust. Soc. Am.* **85**, 868–878.
- Häusler, R., Colburn, S. & Marr, E. (1983) *Acta Otolaryngol. Suppl.* **400**, 1–62.
- Slattery, W.H., III, & Middlebrooks, J.C. (1994) *Hear. Res.* **75**, 38–46.
- Irvine, D.R.F. (1992) in *The Mammalian Auditory Pathway: Neurophysiology*, eds. Popper, A.N. & Fay, R.R. (Springer, New York), pp. 153–231.
- Yu, J.J. & Young, E.D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11780–11786.
- Kuwada, S., Batra, R., Yin, T.C.T., Oliver, D.L., Haberly, L.B. & Stanford, T.R. (1997) *J. Neurosci.* **17**, 7565–7581.
- Li, L. & Kelly, J.B. (1992) *J. Neurosci.* **12**, 4530–4539.
- McAlpine, D., Jiang, D., Shackleton, T.M. & Palmer A.R. (1998) *J. Neurosci.* **18**, 6026–6039.
- Ramachandran, R., Davis, K.A. & May, B.J. (1999) *J. Neurophysiol.* **82**, 152–163.
- Jenkins, W.M. & Masterton, R.B. (1982) *J. Neurophysiol.* **47**, 987–1016.
- Kavanagh, G.L. & Kelly, J.B. (1987) *J. Neurophysiol.* **57**, 1746–1766.
- Middlebrooks, J.C. (2000) in *The New Cognitive Neurosciences*, ed. Gazzaniga, M.S. (MIT Press, Cambridge, MA), pp. 425–436.
- King, A.J., Jiang, Z.D. & Moore, D.R. (1998) *J. Comp. Neurol.* **390**, 342–365.
- Oliver, D.L. & Huerta, M.F. (1992) in *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D.B., Popper, A.N. & Fay, R.R. (Springer, New York), pp. 168–221.
- King, A.J. & Hutchings, M.E. (1987) *J. Neurophysiol.* **57**, 596–624.
- King, A.J. & Palmer, A.R. (1983) *J. Physiol.* **342**, 361–381.
- Middlebrooks, J.C. & Knudsen, E.I. (1984) *J. Neurosci.* **4**, 2621–2634.
- Yin, T.C.T., Hirsch, J.A. & Chan, J.C.K. (1985) *J. Neurophysiol.* **53**, 746–758.
- Stein, B.E. & Meredith, M.A. (1993) *The Merging of the Senses* (MIT Press, Cambridge, MA).
- Hartline, P.H., Pandey Vimal, R.L., King, A.J., Kurylo, D.D. & Northmore, D.P.M. (1995) *Exp. Brain Res.* **104**, 402–408.
- Jay, M.F. & Sparks, D.L. (1984) *Nature (London)* **309**, 345–347.
- Populin, L.C. & Yin, T.C.T. (1998) in *Psychophysical and Physiological Advances in Hearing*, eds. Palmer, A.R., Rees, A., Summerfield, A.Q. & Meddis, R. (Whurr, London), pp. 441–448.
- Middlebrooks, J.C. (1999) *J. Acoust. Soc. Am.* **106**, 1480–1492.

27. Carlile, S. & Pralong, D. (1994) *J. Acoust. Soc. Am.* **95**, 3445–3459.
28. Clifton, R.K., Gwiazda, J., Bauer, J. A., Clarkson, M.G. & Held, R.M. (1988) *Dev. Psychol.* **24**, 477–483.
29. Moore, D.R. & Irvine, D.R.F. (1979) *Acta Otolaryngol.* **87**, 434–440.
30. Werner, L.A. & Gray, L. (1998) in *Development of the Auditory System*, eds. Rubel, E.W., Popper, A.N. & Fay, R.R. (Springer, New York), pp. 12–79.
31. Middlebrooks, J.C. (1999) *J. Acoust. Soc. Am.* **106**, 1493–1510.
32. Wenzel, E.M., Arruda, M., Kistler, D.J. & Wightman, F.L. (1993) *J. Acoust. Soc. Am.* **94**, 111–123.
33. Knudsen, E.I. (1999) *J. Comp. Physiol. A* **185**, 305–321.
34. King, A.J. (1999) *BioEssays* **21**, 900–911.
35. Knudsen, E.I. & Brainard, M.S. (1991) *Science* **253**, 85–87.
36. King, A.J., Hutchings, M.E., Moore, D.R. & Blakemore, C. (1988) *Nature (London)* **332**, 73–76.
37. King, A.J., Schnupp, J.W.H. & Thompson, I.D. (1998) *J. Neurosci.* **18**, 9394–9408.
38. Schnupp, J.W.H. & King A.J. (1997) *J. Neurophysiol.* **78**, 2717–2731.
39. King, A.J. & Carlile, S. (1993) *Exp. Brain Res.* **94**, 444–455.
40. King, A.J. & Parsons, C.H. (1999) *Eur. J. Neurosci.* **11**, 3945–3956.
41. Rauschecker, J.P. & Kniepert, U. (1994) *Eur. J. Neurosci.* **6**, 149–160.
42. Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S.A. & Neville, H.J. (1999) *Nature (London)* **400**, 162–166.
43. Schnupp, J.W.H., King, A.J. & Carlile, S. (1998) *J. Neurophysiol.* **79**, 1053–1069.
44. Parsons, C.H., Lanyon, R.G., Schnupp, J.W.H. & King, A.J. (1999) *J. Neurophysiol.* **82**, 2294–2309.
45. Knudsen, E.I. (1985) *J. Neurosci.* **5**, 3094–3109.
46. Mogdans, J. & Knudsen, E.I. (1992) *J. Neurosci.* **12**, 3473–3484.
47. Hogan, S.C., Stratford, K.J. & Moore, D.R. (1997) *Br. Med. J.* **314**, 350–353.
48. Bauer, R.W., Matuzsa, J.L., Blackmer, R.F. & Glucksberg, S. (1966) *J. Acoust. Soc. Am.* **40**, 441–444.
49. Florentine, M. (1976) *J. Am. Audiol. Soc.* **1**, 243–251.
50. Hofman, P.M., Van Riswick, J.G.A. & Van Opstal, A.J. (1998) *Nat. Neurosci.* **1**, 417–421.
51. McPartland, J.L., Culling, J.F. & Moore, D.R. (1997) *Hear. Res.* **113**, 165–172.
52. Knudsen, E.I., Esterly, S.D. & Knudsen, P.F. (1984) *J. Neurosci.* **4**, 1001–1011.
53. Wilmington, D., Gray, L. & Jahrsdoerfer, R. (1994) *Hear. Res.* **74**, 99–114.
54. Moore, D.R., Hine, J.E., Jiang, Z.D., Matsuda, H., Parsons, C.H. & King, A.J. (1999) *J. Neurosci.* **19**, 8704–8711.
55. Moore, D.R., Hutchings, M.E. & Meyer, S.E. (1991) *Audiology* **30**, 91–101.
56. Pillsbury, H.C., Grose, J.H. & Hall, J.W., III (1991) *Arch. Otolaryngol. Head Neck Surg.* **117**, 718–723.
57. Hall, J.W., III, Grose, J.H. & Pillsbury, H.C. (1995) *Arch. Otolaryngol. Head Neck Surg.* **121**, 847–852.
58. Hogan, S.C., Meyer, S.E. & Moore, D.R. (1996) *Audiol. Neurootol.* **1**, 104–111.

SPATIAL PROCESSING IN THE AUDITORY CORTEX OF THE MACAQUE MONKEY

Gregg H. Recanzone*

Center for Neuroscience and Section of Neurobiology, Physiology, and Behavior, University of California, 1544 Newton Court, Davis, CA 95616

The patterns of cortico-cortical and cortico-thalamic connections of auditory cortical areas in the rhesus monkey have led to the hypothesis that acoustic information is processed in series and in parallel in the primate auditory cortex. Recent physiological experiments in the behaving monkey indicate that the response properties of neurons in different cortical areas are both functionally distinct from each other, which is indicative of parallel processing, and functionally similar to each other, which is indicative of serial processing. Thus, auditory cortical processing may be similar to the serial and parallel “what” and “where” processing by the primate visual cortex. If “where” information is serially processed in the primate auditory cortex, neurons in cortical areas along this pathway should have progressively better spatial tuning properties. This prediction is supported by recent experiments that have shown that neurons in the caudomedial field have better spatial tuning properties than neurons in the primary auditory cortex. Neurons in the caudomedial field are also better than primary auditory cortex neurons at predicting the sound localization ability across different stimulus frequencies and bandwidths in both azimuth and elevation. These data support the hypothesis that the primate auditory cortex processes acoustic information in a serial and parallel manner and suggest that this may be a general cortical mechanism for sensory perception.

One of the fundamental tasks of the auditory system is to determine the spatial location of acoustic stimuli. In contrast to the visual and somatosensory systems, the auditory periphery cannot encode stimulus location, but can only encode the presence of particular stimulus frequencies in the input. The central nervous system therefore must compute the spatial location of a stimulus by integrating the responses of many individual sensory receptors.

There are three main cues that can be used to compute the spatial location of an acoustic stimulus: interaural intensity, interaural time or phase, and differences in the stimulus spectrum at the tympanic membrane (1). The binaural cues are critical for localization in azimuth, but are much less effective for localization in elevation because the ears of most mammals are located symmetrically on the head. However, reflections of the acoustic signal by the torso, head, pinna, and ear canal create spectral peaks and notches that vary with stimulus elevation (2, 3). Although the physical cues that could provide the necessary information to localize sounds are well defined, how the nervous system uses these cues to calculate the spatial location of acoustic stimuli is far from being resolved. There are several stations along the ascending auditory pathway in mammals that integrate the spatial cues necessary for the localization of sounds, including the superior olivary complex (4), the inferior colliculus (5–7), and the thalamus (8–10). The spatial tuning of the majority of auditory cortical neurons is very broad, commonly over 90° for a half-maximal response (11–16). In contrast, primates can detect changes in sound location as small as a few degrees or less (17–22). This finding may appear to indicate that auditory cortex is not necessary for this perception, but auditory cortical lesions produce clear deficits in sound localization performance in cats (23), ferrets (24), New World monkeys (25), Old World monkeys (26), and humans (27). Thus, a key question is how the broad spatially tuned neurons in the auditory cortex processes acoustic information to ultimately result in the perception of acoustic space.

The auditory cortex of the primate can be anatomically subdivided into several “core”, “belt,” and “parabelt” cortical areas based on cytoarchitecture, cortico-cortical connections, and thalamo-cortical connections (see refs. 28 and 29). It has been speculated that these multiple auditory cortical areas process acoustic information in both a serial and parallel manner (28) similar to visual cortical processing of “what” and “where” information (30, 31). While the available anatomical data are consistent with this hypothesis, there are relatively few electrophysiological studies in the monkey to either support or refute this idea. Merzenich and Brugge (32) were the first to describe the physiological properties of the macaque primary auditory cortex (AI) and the rostral field (R) in the core area, and the caudomedial (CM) field and lateral field (L) of the “belt” area (32). They found that AI and R neurons had sharper frequency tuning than those in CM based on the multiple-unit responses in the anesthetized animal. Subsequent studies (33, 34) support these initial observations. More recent studies indicate that neurons in the L of the belt area respond better to spectrally complex stimuli, including vocalizations (35), which suggests that the L is processing “what” information. In contrast, caudal and medial fields have been proposed to process “where” information. Neurons in CM have broad frequency tuning and the responses of CM neurons to tone stimuli depend on an intact AI (36). These limited physiological data are consistent with serial processing of acoustic information from the core to the belt auditory cortical areas, and this relatively new hypothesis currently is being rigorously tested in several laboratories.

Neuronal Activity as a Function of Stimulus Frequency and Intensity

Previous electrophysiological studies in the primate auditory cortex have largely been done in anesthetized animals. However, the activity of neurons in the primate auditory cortex can either increase or decrease depending on whether the monkey is attending to the stimulus, not attending to the stimulus, or is anesthetized (11, 37, 38). To define the frequency and intensity responses of primate cortical neurons in the attended state, single neuron responses were recorded in monkeys while they performed a sound localization task (39). In this experiment,

*E-mail: ghrecanzone@ucdavis.edu.

This paper was presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: AI, primary auditory cortex; CM, caudomedial; FRA, frequency response area; CF, characteristic frequency; R, rostral field; L, lateral field; PSTH, poststimulus time histogram.

tone stimuli at 31 different frequencies (2- to 5-octave range) and 16 different intensities (90-dB range) were presented from a speaker located directly opposite to the contralateral ear. Fig. 1 shows representative frequency response areas (FRAs) measured across three different auditory cortical areas in a representative monkey. The normalized firing rate for each stimulus is indicated by the color, with red regions corresponding to the stimuli that elicited the greatest activity and blue showing stimuli that elicited activity significantly greater than the spontaneous rate but less than 25% of the peak activity. The frequency range tested was adjusted for each neuron, as the frequency tuning could be quite different between neurons in different cortical areas (Fig. 1 Upper Right). These experiments demonstrated that AI neurons in the behaving monkey had relatively sharp frequency tuning (e.g., Fig. 1 B, C, E, and F). In contrast, neurons in CM generally had broader frequency tuning (Fig. 1 D, G, and H), even for neurons with similar characteristic frequencies (CFs), defined as the frequency that elicited a response at the lowest intensity (Fig. 1 C vs. G). There also was a shift in CF when crossing the border between different cortical areas, for example from L to AI (Fig. 1 A vs. B) or between AI and CM (Fig. 1 C vs. D and Fig. 1 F vs. G).

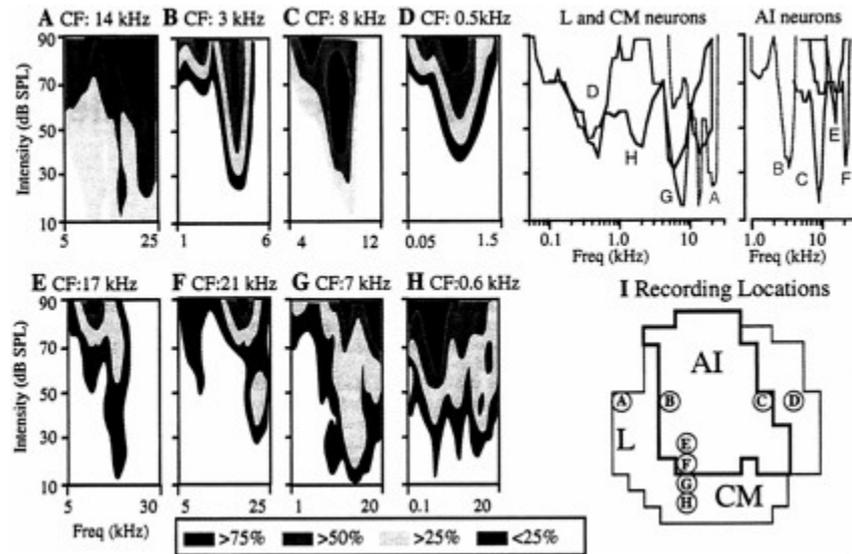


Fig. 1. Frequency response areas of single auditory cortical neurons. Responses were recorded to 50-ms tone stimuli (3-ms rise/fall) presented at 16 different intensity levels [10- to 90-dB sound pressure level (SPL)] at 31 different frequencies spanning 2–5 octaves from a free-field speaker located directly opposite to the contralateral ear. The color corresponds to the percent of the maximum response recorded in that neuron. White areas correspond to areas where the activity was not significantly greater than the spontaneous rate. Each FRA shows the response of a single neuron from the cortical location shown in *I*. The frequency range was customized for each neuron and therefore will vary between panels. The 25% contour (50% for the neuron shown in *A*) is reproduced on the same frequency axis to allow comparisons of the frequency bandwidth across neurons (Upper Right), and the CF is given above each FRA. (*I*) Dorsal view of the recording locations for each neuron. The heavy line shows the physiological boundaries of AI. Thin lines show the region investigated in the study. Circled letters correspond to the different panels shown in the figure. Note the differences in frequency tuning between neurons in AI and other cortical fields. Adapted from ref. 39.

Fig. 2 shows representative FRAs recorded from neurons in a second monkey. Neurons in R showed similar tuning functions as AI neurons (compare Fig. 2 A-G to H-K, M, and O). Again, the AI and CM border was easily identified in this monkey by the change in the frequency tuning and the CF (Fig. 2 K-P).

These observations are consistent with those described in the anesthetized monkey and indicate that different auditory cortical areas have distinct functional properties using simple tone stimuli. Statistical analysis confirmed that CM neurons had the broadest frequency tuning of all fields examined, and neurons in R had the narrowest frequency tuning (39). The ability to integrate information across a broad frequency range would likely improve spatial processing, as binaural and spectral cues across different frequencies could be used, and broadband stimuli are more easily localized than narrow band stimuli (see below). The broad frequency tuning of neurons in CM would make them ideally suited to integrate information across frequencies, consistent with the hypothesis that AI and CM form part of a serial “where” processing stream of auditory information (28).

Neuronal Activity as a Function of Stimulus Location

The hypothesis that AI and CM neurons process auditory spatial information in series predicts that the spatial response properties of these neurons should improve between AI and CM. To address this issue, the responses of neurons in these areas were measured while the monkey performed a sound localization task, and the neuronal activity was compared with the monkey’s sound localization performance (16). To determine sound localization thresholds, the monkey depressed a lever to initiate a trial, and several stimuli were presented from directly in front of the monkey. At some random time the stimulus changed location in either azimuth or elevation. When the monkey detected this change it released the lever and received a reward. The sound localization threshold was defined as the distance between locations necessary for the monkey to detect a difference on half

of the trials. These thresholds are shown for two monkeys in Fig. 3. The filled bars show the thresholds measured in azimuth and the open bars show thresholds measured in elevation for tone stimuli of different frequencies (*Left*) or noise stimuli with different spectral content (*Right*). Across these different stimuli, the thresholds for localization in azimuth were lower than those for localization in elevation. This difference was greatest for tone stimuli, where in most cases the elevation thresholds could not be measured because 30° was the maximum change in location tested. For noise stimuli, there was a progressive improvement in elevation thresholds as the stimulus contained higher frequency components. The worst thresholds were noted for stimuli containing 750–1,500 Hz, improving for 3,000–6,000 Hz, and 5,000–10,000 Hz, and the lowest thresholds were noted when the stimulus was a broadband noise containing all of those frequencies. There was no such obvious trend as a function of the tone stimulus frequency.

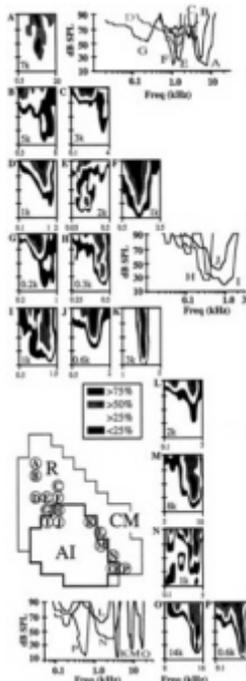


Fig. 2. FRAs recorded in the three different auditory cortical areas in a second monkey. (A-G) FRAs from single neurons recorded in R. (H-J) FRAs recorded at the rostral border of AI. (K, M, and O) Neurons recorded at the medial border of AI. (L, N, and P) Neurons recorded in CM near the AI-CM border. The characteristic frequency is shown within each FRA. Other conventions are as in Fig. 1. Adapted from ref. 39.

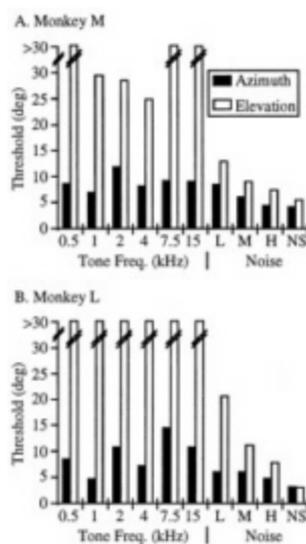


Fig. 3. Sound localization thresholds across stimulus frequencies and bandwidths. Thresholds are shown for localization in azimuth (solid bars) and elevation (open bars). Thresholds could not be defined if they were greater than 30° (broken lines). Noise stimuli consisted of 1-octave band-passed noise (L: 750–1,500 Hz; M: 3,000–6,000 Hz; H: 5,000–10,000 Hz) and broadband noise (NS). Adapted from ref. 16.

The activity of single neurons also was recorded in these monkeys while they performed a similar task. Each neuron was tested with two stimuli on randomly interleaved trials. One stimulus was a tone near the characteristic frequency of the neuron and the other was a noise stimulus that included the CF. Both of these stimuli usually elicited a robust response from the neuron under study. A typical example from an AI neuron is

shown in Fig. 4. To the left are poststimulus time histograms (PSTHs) taken over 10 trials in which either a tone (Fig. 4A) or band-passed noise (Fig. 4B) was presented from one of 17 different locations in front of the monkey. Stimuli were positioned straight ahead and at 15° and 30° eccentricity along the horizontal, vertical, and both oblique axes. Fig. 4 shows the PSTHs at their relative locations in this region of frontal space. This neuron had a more robust response when the stimuli were presented to the right of the midline (in contralateral space), compared with when the stimuli were presented to the left of the midline. However, there was little difference in activity as a function of the elevation of the stimulus. This can be most readily appreciated by comparing the middle row of PSTHs (azimuth tuning) to the middle column of PSTHs (elevation tuning). The three-dimensional reconstruction of these responses are shown to the right of each plot in Fig. 4. These plots were normalized to the peak activity of that neuron measured across all locations for both stimuli, with the response shown in the z axis as a function of the stimulus azimuth and elevation. The response contour for noise stimuli had a greater slope than the response contour for the tone stimuli, indicating that this neuron was more sensitive to the location of noise stimuli compared with tones.

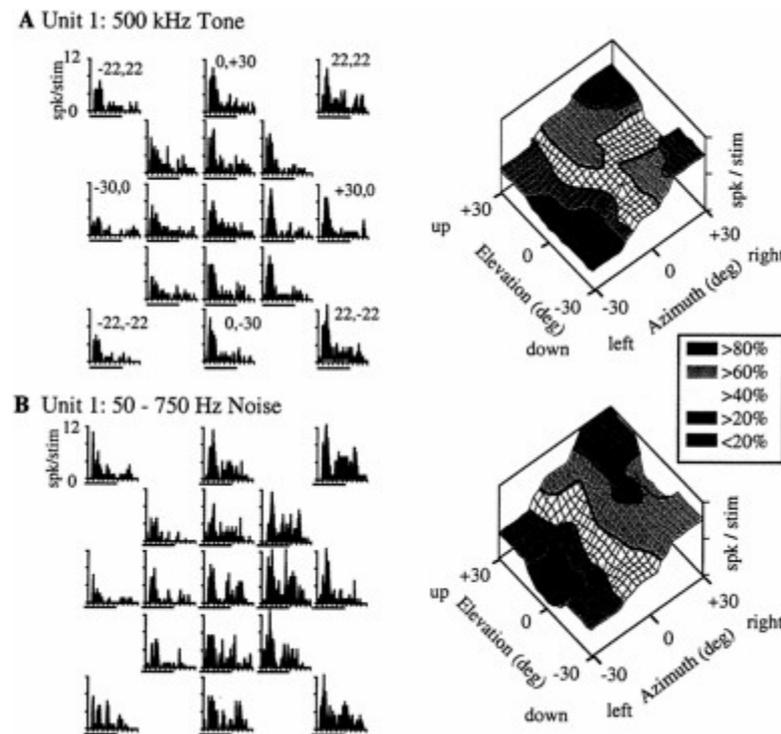


Fig. 4. Spatial response profiles of an AI neuron. PSTHs are shown in their relative position from the monkey's perspective (rightward PSTHs correspond to stimuli presented to the right of midline). Numbers above the most eccentric PSTHs correspond to the location in degrees (azimuth, elevation). Each PSTH shows the responses over 10 trials. Tone (A) and noise (B) stimuli were presented on randomly interleaved trials. In the color-coded three-dimensional plots the response was normalized by the maximum response recorded for that neuron to any of the 17 locations using either the tone or noise stimulus. The magnitude of the response at each azimuth and elevation is shown by the height of the contour. Heavy lines show regions with the same activity (iso-response contours).

An example from a CM neuron is shown in Fig. 5. In this case, the neuron responded better to noise than to tones. Further, the response to noise was more strongly modulated by the stimulus location, illustrated by the greater slope of the surface contour. Finally, there was a difference in the spatial preferences of this neuron depending on the stimulus. When tone stimuli were presented (Fig. 5A), there was essentially no modulation of the response as a function of the stimulus elevation, shown as the iso-intensity contours (heavy black lines) being roughly parallel to the elevation axis. This can best be seen for stimuli at the midline (0° azimuth), where the neuronal response varied very little over 60° differences in elevation when tones were presented. In contrast, the response along the midline to noise stimuli (Fig. 5B) was greatest for upward elevations and smallest at the lowest stimulus elevation.

The results from both monkeys indicated that although most neurons responded to all stimulus locations, i.e., they were very broadly tuned, the main features of the neuronal responses were consistent with the behavioral ability to localize sounds. Localization in elevation was very poor for tone stimuli, and few neurons (<10%) were encountered that had changes in their response as a function of the elevation of tone stimuli. In contrast, localization in elevation of noise stimuli containing high-frequency components was much better than for tone stimuli, and more neurons were encountered that were sensitive

to the elevation of these noise stimuli ($\approx 40\%$). Secondly, there was a greater rate of change in the response as a function of the stimulus azimuth for noise stimuli compared with tone stimuli. Finally, the highest percentage of neurons were sensitive to the location of broadband noise ($\approx 55\%$ in azimuth and $\approx 30\%$ in elevation for AI neurons and $\approx 80\%$ in azimuth and $\approx 30\%$ in elevation for CM neurons), which showed the lowest behavioral thresholds of all stimuli tested. These general observations suggest that the firing rate of single neurons could contain sufficient information for the monkey to localize these different types of stimuli.

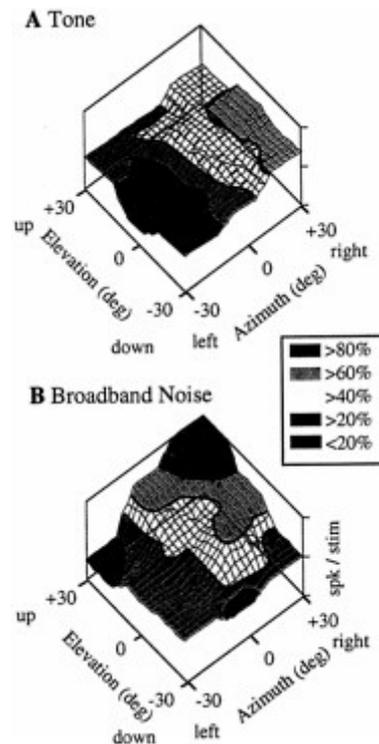


Fig. 5. Three-dimensional reconstructions of spatial responses from a representative CM neuron. Conventions are as in Fig. 4. This neuron had a lower response to tone stimuli, and a more shallow response as a function of stimulus location (A). The response to noise stimuli (B) showed greater modulation as a function of stimulus location, and the slope of this response contour was not aligned with either the elevation or azimuth axis. This indicates that this neuron contained information for both the azimuth and elevation of the stimulus.

Correlations Between Neural Activity and Sound Localization

These qualitative impressions were verified by directly comparing the neuronal and behavioral data. Fig. 6 shows the firing rate as a function of stimulus azimuth for a single AI neuron (A) and a single CM neuron (B). The task that was used to define thresholds (Fig. 3) required the monkey to detect a change in the location of the stimulus from directly ahead. If the monkey had access to the information provided by only one neuron, then significant differences in activity from when the stimulus was presented directly in front of the monkey would be a reliable signal that the stimulus had changed location. The predicted threshold for each neuron was defined as the distance that the stimulus would have to move for the activity to change by one standard deviation from when the stimulus was straight ahead (dashed lines of Fig. 6). This predicted threshold would be large if the spatial tuning of the neuron was relatively poor (slopes of the line near zero), and the predicted threshold would be small if the response of the neuron was strongly modulated by stimulus position. The predicted threshold was compared with the behavior by taking the ratio of the predicted threshold divided by the measured threshold. This ratio was less than one if the neuron predicted a smaller threshold than was observed, one if the neuron and the behavior were the same, and greater than one if the behavioral threshold was smaller than the neuronal prediction. If the neuronal responses reflect the sound localization ability, stimuli that the monkey had difficulty in localizing should elicit poor spatial resolution in most neurons (and therefore predict high thresholds for a ratio near 1.0), while stimuli that the monkey could easily localize should elicit sharp spatial resolution in most neurons (and therefore predicted low thresholds for a ratio near 1.0). The distribution of this ratio for 353 AI neurons and 118 CM neurons is shown in the middle panel of Fig. 6. For both AI and CM, while most neurons predicted thresholds greater than those observed behaviorally, many neurons did predict thresholds consistent with the behavior. Further, CM neurons were better able to predict the behavior than AI neurons ($P < 0.05$) as indicated by more neurons having ratios close to 1.0 (compare the middle and right panels of Fig. 6 A and B).

The ability of some neurons to predict behavior indicates that neurons in these areas could provide valuable information to the monkey about the spatial location of the stimulus. However, there was a wide variation in threshold ratios, meaning that many cells performed better or worse than the monkey. Because all neurons responded to these stimuli they were presumably conveying some information to the monkey. One possibility is that pooling the responses of all neurons would enhance the ability to predict the behavior. Alternatively, it may be that pooling the responses of all neurons would cause a degradation of the ability of the population to predict the behavior caused by the neurons that showed poor spatial sensitivity.

The results of an analysis of pooling neurons is shown in Fig. 7 where the mean and standard deviation across all comparisons (tone and noise stimuli for azimuth and elevation, 21 comparisons total) are shown for two populations of pooled neurons in each cortical area. Open bars show the results when all neurons tested in each cortical area were pooled. The neuronal predictions of the behavior for both AI and CM neurons were significantly worse than the measured behavior. A second level of analysis pooled the responses based on their spatial tuning. Significant spatial tuning was defined as a statistically significant correlation of the response as a function of stimulus location in at least one direction (azimuth or elevation) for at least one tested stimulus (tone or noise). The closed bars of Fig. 7 show that there was an improvement in the ability to predict the behavior by pooling the responses of only these spatially sensitive neurons. For AI neurons, the improvement still resulted in predictions that were significantly worse than the measured behavior. For CM neurons, however, the predictions based on the pooled spatially sensitive neurons were not different from the behavioral thresholds measured in the monkey. This result indicates that relatively small populations of neurons in CM contained sufficient information for the monkey to perform the task.

These results are consistent with serial processing of spatial information from AI to CM in the primate auditory cortex.

Neurons in CM showed better spatial tuning than AI neurons, and the ability to predict the behavior by all of the measured CM neurons was not significantly different from the predictions by only the spatially sensitive AI neurons. This is expected if the CM neurons were selectively activated by the spatially tuned AI neurons, ultimately leading to an enhanced representation of acoustic space during this serial processing. In support of this idea is the finding that CM neurons receive inputs from broad regions of AI that span much of the frequency representation (33, 36).

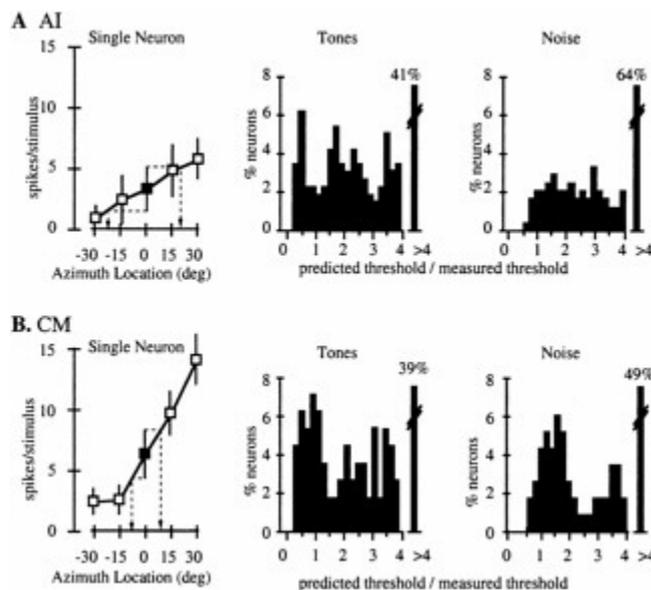


Fig. 6. Predictions of behavioral performance by single neurons. The mean and standard deviation of the response of a single neuron as a function of the stimulus azimuth (0° elevation) are shown for an AI neuron (A) and a CM neuron (B). (Left) notes the response from the speaker located directly in front of the monkey. The ability of the neuron to predict the behavior was calculated as the distance in azimuth that corresponded to one standard deviation from the mean response at 0° (dashed lines). This prediction was tested against the behavior by dividing the predicted threshold by the measured threshold. (Center) The frequency distribution of this ratio when predicting thresholds in azimuth for tone stimuli measured across 353 AI neurons (A) and 118 CM neurons (B). Neurons that had a prediction greater than four times the measured threshold are shown in the right most bin. Ratios of 1.0 correspond to perfect predictions. (Right) The ratios when predicting thresholds in azimuth for noise stimuli. Adapted from ref. 16.

These results raise several obvious questions. The first is which other cortical areas also process spatial information. The experiments to date have concentrated on AI and CM, but it remains to be seen how neurons in other cortical areas also participate in this perception. It is likely that other cortical areas also will have spatially tuned neurons, as the “parallel” nature of information processing is almost certainly not strictly maintained. It is more likely that neurons across cortical areas process both “what” and “where” information to differing degrees to aid in ultimately “binding” these features to give rise to the percept of a real-world object.

A second question is how this information is used by the monkey. Although both AI and CM are likely to be necessary for sound location perception, it is also unlikely that either is sufficient for this percept in the primate. The most likely scenario is that these neurons form one link in the serial processing of spatial information that will be further processed in other auditory cortical fields, as well as parietal (e.g., ref. 40) and/or frontal cortical areas (41). The inputs from CM are likely

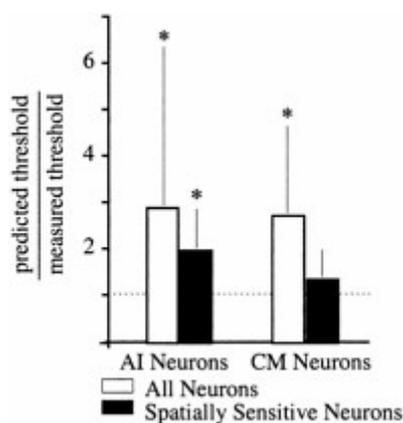


Fig. 7. Mean and standard deviation for the predicted/measured ratio pooled across either all neurons measured in that cortical area (open bars) or restricted to only the neurons in that cortical area that had significant correlation between the neuronal activity and the spatial location for at least one stimulus (closed bars). Each bar represents the mean of the azimuth and elevation predictions for tone and noise stimuli (21 ratios total). Dashed line is through 1.0 (perfect prediction). Only the pooled spatially sensitive CM neurons had a ratio that was not significantly different from the behavior. Adapted from ref. 16.

candidates to contribute to the creation of multimodal spatial perception in the parietal lobe (40).

In summary, the available physiological evidence is supportive of the hypothesis that spatial location is processed in series between AI and CM. It remains to be seen how the outputs of CM are further processed, and how this processing results in the perception of acoustic space. Similarly, other features of the acoustic stimulus may be preferentially processed in other cortical areas, for example in the L fields in the belt and parabelt areas. Finally, the role of the cortical areas in the core region, particularly AI and R, is still unclear. It may be that both areas process all types of information in parallel, or there may be a subdivision of feature processing at this initial cortical level. Nonetheless, these experiments on the cortical mechanisms of sound localization indicate that broadly tuned neurons can in fact provide information necessary to perform perceptual discriminations at a much finer resolution than the bandwidths of the neuronal tuning functions would suggest. This type of information processing may be a general mechanism by which the activity of neurons in the cerebral cortex leads to perception across sensory modalities (42–44).

I thank E.A. Disbrow, L.A. Krubitzer, J.A. Langston, M.L. Phan, T.K. Su, and T.M. Woods for helpful comments on earlier versions of this manuscript, and D.C. Guard, M.L. Phan and T.K. Su for their participation in the described experiments. Funding was provided by National Institutes of Health Grant DC-02371, the Klingenstein Foundation, and the Sloan Foundation.

1. Middlebrooks, J.C. & Green, D.M. (1991) *Annu. Rev. Psychol.* **42**, 135–159.
2. Wightman, F.L. & Kistler, D.J. (1989) *J. Acoust. Soc. Am.* **85**, 868–878.
3. Pralong, D. & Carlile, S. (1994) *J. Acoust. Soc. Am.* **95**, 3435–3444.
4. Joris, P.X. & Yin, T.C. (1995) *J. Neurophysiol.* **73**, 1043–1062.
5. Yin, T.C. & Chan, J.C. (1990) *J. Neurophysiol.* **64**, 465–488.
6. Kuwada, S., Batra, R., Yin, T.C., Oliver, D.L., Haberly, L.B. & Stanford, T.R. (1997) *J. Neurosci.* **17**, 7565–7581.
7. Litovsky, R.Y. & Yin, T.C. (1998) *J. Neurophysiol.* **80**, 1285–1301.
8. Clarey, J.C., Barone, P. & Imig, T.J. (1994) *J. Neurophysiol.* **72**, 2384–2405.
9. Barone, P., Clarey, J.C., Irons, W.A. & Imig, T.J. (1996) *J. Neurophysiol.* **75**, 1206–1220.
10. Imig, T.J., Poirier, P., Irons, W.A. & Samson, F.K. (1997) *J. Neurophysiol.* **78**, 2754–2771.
11. Benson, D.A., Hienz, R.D. & Goldstein, M.H. Jr. (1981) *Brain Res.* **219**, 249–267.
12. Middlebrooks, J.C. & Pettigrew, J.D. (1981) *J. Neurosci.* **1**, 107–120.
13. Imig, T.J., Irons, W.A. & Samson, F.R. (1990) *J. Neurophysiol.* **63**, 1448–1466.
14. Rajan, R., Aitkin, L.M., Irvine, D.R.F. & McKay, J. (1990) *J. Neurophysiol.* **64**, 872–887.
15. Middlebrooks, J.C., Xu, L., Eddins, A.C. & Green, D.M. (1998) *J. Neurophysiol.* **80**, 863–881.
16. Recanzone, G.H., Guard, D.C., Phan, M.L. & Su, T.K. (2000) *J. Neurophysiol.* **83**, 2723–2739.
17. Alshuler, M.W. & Comalli, P.E. (1975) *J. Aud. Res.* **15**, 262–265.
18. Brown, C.H., Beecher, M.D., Moody, D.B. & Stebbins, W.C. (1978) *Science* **201**, 753–754.
19. Brown C.H., Beecher M.D., Moody D.B. & Stebbins W.C. (1980) *J. Acoust. Soc. Am.* **68**, 127–132.
20. Brown, C.H., Schessler, T., Moody, D. & Stebbins, W. (1982) *J. Acoust. Soc. Am.* **72**, 1804–1811.
21. Perrott, D.R. & Saberi, K. (1990) *J. Acoust. Soc. Am.* **87**, 1728–1731.
22. Recanzone, G.H., Makhamra, S.D.D.R. & Guard, D.C. (1998) *J. Acoust. Soc. Am.* **103**, 1085–1097.
23. Jenkins, W.M. & Merzenich, M.M. (1984) *J. Neurophysiol.* **52**, 819–847.
24. Kavanagh, G.L. & Kelly, J.B. (1987) *J. Neurophysiol.* **57**, 1746–1766.
25. Thompson, G.C. & Cortez, A.M. (1983) *Behav. Brain Res.* **8**, 211–216.
26. Heffner, H.E. & Heffner, R.S. (1990) *J. Neurophysiol.* **64**, 915–931.
27. Sanchez-Longo, L.P. & Forster, F.M. (1958) *Neurology* **8**, 119–125.
28. Rauschecker, J.P. (1998) *Audiol. Neuro-Otol.* **3**, 86–103.
29. Kaas, J.H., Hackett, T.A. & Tramo, M.J. (1999) *Curr. Opin. Neurobiol.* **9**, 164–170.
30. Ungerleider, L.G. & Mishkin, M. (1982) in *Analysis of Visual Behavior*, eds Ingle, D.J., Goodale, M.A. & Mansfield, J.W. (MIT Press, Cambridge, MA), pp. 549–556.
31. Ungerleider, L.G. & Haxby, J.V. (1994) *Curr. Opin. Neurobiol.* **4**, 157–165.
32. Merzenich, M.M. & Brugge, J.F. (1973) *Brain Res.* **50**, 275–296.
33. Morel, A., Garraghty, P.E. & Kaas, J.H. (1993) *J. Comp. Neurol.* **335**, 437–459.
34. Kosaki, H., Hashikawa, T., He, J. & Jones, E.G. (1997) *J. Comp. Neurol.* **386**, 304–316.
35. Rauschecker, J.P., Tian, B. & Hauser, M. (1995) *Science* **268**, 111–114.
36. Rauschecker, J.P., Tian, B., Pons, T. & Mishkin, M. (1997) *J. Comp. Neurol.* **382**, 89–103.
37. Miller, J.M., Sutton, D., Pflingst, B., Ryan, A., Beaton, R. & Gourevitch, G. (1972) *Science* **111**, 449–451.
38. Miller, J.M., Dobie, R.A., Pflingst, B.E. & Hienz, R.D. (1980) *Am. J. Otolaryngol.* **1**, 119–130.
39. Recanzone, G.H., Guard, D.C. & Phan, M.L. (2000) *J. Neurophysiol.* **83**, 2315–2331.
40. Mazzoni, P., Bracewell, R.M., Barash, S. & Andersen, R.A. (1996) *J. Neurophysiol.* **75**, 1233–1241.
41. Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S. & Rauschecker, J.P. (1999) *Nat. Neurosci.* **2**, 1131–1136.
42. Bradley, A., Skottun, B.C., Ohzawa, I., Sclar, G. & Freeman, R.D. (1987) *J. Neurophysiol.* **57**, 755–772.
43. Vogels, R. & Orban, G.A. (1990) *J. Neurosci.* **10**, 3543–3558.
44. Prince, S.J., Pointon, A.D., Cumming, B.G. & Parker, A.J. (2000) *J. Neurosci.* **20**, 3387–3400.

SONG SELECTIVITY AND SENSORIMOTOR SIGNALS IN VOCAL LEARNING AND PRODUCTION

Michele M.Solis*, Michael S.Brainard†, Neal A.Hessler†, and Allison J.Doupe†‡

†Keck Center for Integrative Neuroscience, and Departments of Physiology and Psychiatry, University of California, San Francisco, CA 94143; and *Department of Psychology, University of Washington, Seattle, WA 98195

Bird song, like human speech, is a learned vocal behavior that requires auditory feedback. Both as juveniles, while they learn to sing, and as adults, songbirds use auditory feedback to compare their own vocalizations with an internal model of a target song. Here we describe experiments that explore a role for the songbird anterior forebrain pathway (AFP), a basal ganglia-forebrain circuit, in evaluating song feedback and modifying vocal output. First, neural recordings in anesthetized, juvenile birds show that single AFP neurons are specialized to process the song stimuli that are compared during sensorimotor learning. AFP neurons are tuned to both the bird's own song and the tutor song, even when these stimuli are manipulated to be very different from each other. Second, behavioral experiments in adult birds demonstrate that lesions to the AFP block the deterioration of song that normally follows deafening. This observation suggests that deafening results in an instructive signal, indicating a mismatch between feedback and the internal song model, and that the AFP is involved in generating or transmitting this instructive signal. Finally, neural recordings from behaving birds reveal robust singing-related activity in the AFP. This activity is likely to originate from premotor areas and could be modulated by auditory feedback of the bird's own voice. One possibility is that this activity represents an efference copy, predicting the sensory consequences of motor commands. Overall, these studies illustrate that sensory and motor processes are highly interrelated in this circuit devoted to vocal learning, as is true for brain areas involved in speech.

Human speech and bird song share numerous features (1). Both are complex acoustic sequences, generated by coordinated actions of the vocal apparatus and the muscles of respiration. Most importantly, both speech and song are learned and are strongly influenced by hearing in early life and in adulthood: neither birds nor humans learn to vocalize normally in the absence of hearing, and as adults, both show deterioration of vocal output after hearing loss (2–6). Songbirds thus provide a promising model system for elucidating general neural mechanisms involved in vocal learning, including how the brain evaluates auditory feedback and uses it to modify vocal output.

Experiments to investigate the neural basis of vocal learning in songbirds are aided by a wealth of information about the behavioral time course of learning (7–9) and its dependence on hearing (4, 6). Song learning occurs in two stages, called the sensory and sensorimotor phases (Fig. 1A). During the sensory phase, a young bird listens to and memorizes the song of an adult tutor, often the bird's father. This memory is called the template. The sensorimotor phase begins later, when the young bird begins to sing; during sensorimotor learning the juvenile uses auditory feedback to compare its own immature vocalizations (plastic song) to the tutor song template, and gradually refines and adapts its vocal output until it matches the template. Thus, auditory experience of both the tutor song and the bird's own song (BOS) is required during learning. In adulthood, elimination or alteration of auditory feedback of BOS induces gradual deterioration of adult song structure (5, 10). These behavioral observations suggest that there must be neural circuitry involved in memorization and evaluation of song. Specifically, there must be mechanisms that compare auditory feedback from vocal output to the internal song template and that generate signals to guide changes in vocal output.

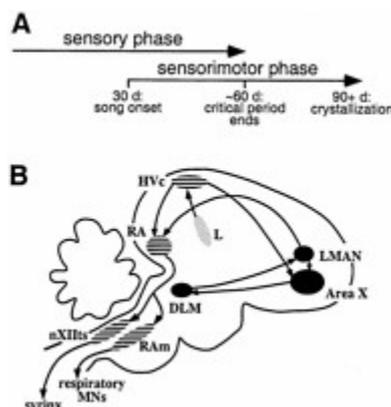


Fig. 1. (A) Song learning occurs in two phases. For zebra finches, the sensory phase ends at ≈ 60 days of age and the sensorimotor phase begins when birds are ≈ 30 days old and continues until they are ≥ 90 days of age; thus the phases of learning overlap in this species. (B) Anatomy of the song system, which consists of two pathways. Motor pathway nuclei are striped, and AFP pathway nuclei are in black. The motor pathway, necessary for normal song production throughout life, includes HVC, the robust nucleus of the archistriatum (RA), and the tracheosyringeal portion of the hypoglossal nucleus (nXIIIts). RA also projects to nuclei involved in control of respiration, such as nucleus retroambiguus (Ram). The AFP comprises Area X (X), the medial nucleus of the dorsolateral thalamus (DLM), and LMAN. The Field L complex and related areas (stippled) provide auditory input to the song system.

One candidate circuit for processing and evaluating these song experiences is the anterior forebrain pathway (AFP), a basal ganglia-forebrain circuit found within a system of interconnected nuclei dedicated to song learning and production (Fig. 1B; ref. 11). The AFP plays a special, but unclear, role during learning. Lesions of the AFP severely disrupt song learning in juveniles, whereas the same lesions do not affect song in normal adults

‡To whom reprint requests should be addressed.

This paper was presented at the National Academy of Sciences colloquium "Auditory Neuroscience: Development, Transduction, and Integration," held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviations: AFP, anterior forebrain pathway; BOS, bird's own song; LMAN, lateral magnocellular nucleus of the anterior neostriatum.

(12–14). The output nucleus of the AFP, the lateral magnocellular nucleus of the anterior neostriatum (LMAN), projects to the motor pathway for song, which is necessary for normal song production throughout life (11). Thus, the AFP is well positioned to influence activity in the motor pathway and could drive changes in vocal output. We review here experiments that implicate AFP function in the sensory and sensorimotor phases of learning, as well as in the sensorimotor processes crucial to song maintenance in adulthood.

Song Learning in Juveniles

As might be expected of neural circuits that may be involved in mediating song learning, neurons in the AFP are responsive to song stimuli. In adult, anesthetized zebra finches, these neurons respond more strongly to BOS than to acoustically similar songs of other zebra finches (conspecific songs; Fig. 2) or BOS played in reverse (15). The properties of these neurons are very similar to those of song-selective neurons first described in the song nucleus HVC (Fig. 1B; refs. 16 and 17). Neurons that are sensitive to the complex spectral and temporal properties of song could be useful for processing song stimuli during learning. Moreover, this song selectivity emerges during the course of song learning: AFP neurons from birds early in the sensory learning phase (30 days of age) respond equally well to all song stimuli, and then over time increase their response to their own song while losing responsiveness to other stimuli (15, 18) (Fig. 2B). There is a striking parallel to this result in human speech development: human infants initially show sensory discrimination of phonemes from all human languages tested, but gradually lose their capacity to accurately discriminate sounds that they are not experiencing, and improve their discrimination of the sounds of the language spoken around them (19–21). In both cases, the initial broad sensitivity endows the young organism with the capacity to learn any language or species-specific song, but this sensitivity then is narrowed and shaped by experience.

Song selectivity develops rapidly, because it is found in the AFP of zebra finches that have completed the sensory phase of learning (60 days of age; Figs. 1A and 2B; ref. 18). At this time zebra finches are also in the middle of the sensorimotor phase and have been producing plastic song for about a month. Thus, experience of either the BOS or the tutor song could have shaped the selectivity of these neurons. Knowing which experience is responsible for selectivity would inform our hypotheses about AFP function during song learning. For example, neurons tuned by BOS experience could provide information about the current state of BOS, whereas those tuned by tutor song could encode the tutor song memory. When we compared the neural responses to BOS and tutor song in 60-day-old birds, we found a range of preferences for one song over the other (Fig. 3A). Many neurons preferred BOS over tutor song, supporting a role for BOS experience in shaping selectivity. A few neurons preferred tutor over BOS, suggesting that they were tuned by tutor song experience. Finally, many neurons responded equally well to both songs. These neurons were clearly selective, because they did not respond as well to conspecific or reversed song stimuli. Thus, such neurons might reflect experiences of both BOS and tutor song.

Two important caveats exist with respect to the apparent shaping of AFP neurons by these two sensory experiences. First, although BOS selectivity initially might seem to reflect the bird's experience of its own song, it is also possible that it actually represents the template. If a bird memorized the tutor song poorly during sensory learning, then modeled its own song after this inaccurate template, BOS selectivity would be a better representation of the template than the tutor song. The question of whether BOS indeed reflects the bird's own vocalizations could be solved if the bird were made to sing something very different from its tutor by a manipulation of its peripheral vocal system. Because the bird would hear the highly abnormal BOS only as a result of its own singing, neurons tuned to the abnormal song would verify that it was the experience of BOS that was critical. Second, neurons tuned to both BOS and tutor song might not reflect the experience of both of these songs, but simply reflect acoustic similarities between these two stimuli. The bird is trying to model its own song after the tutor song, and by 60 days of age, plastic song often resembles the tutor song. This question also could be addressed if the acoustic similarity that normally develops between BOS and tutor song were minimized by inducing juvenile zebra finches to sing abnormal songs (22). If the neurons that respond equally well to BOS and tutor song actually are shaped by the experience of the bird's voice but respond to both stimuli because of acoustic similarities between these songs, then this kind of neuron should not exist in birds with song unlike their tutor song (Fig. 3B). Alternatively, if these neurons reflect independent contributions of both BOS and tutor song experience to selectivity, then they should persist in birds with song unlike their tutor song, perhaps as separate neural populations (Fig. 3C).

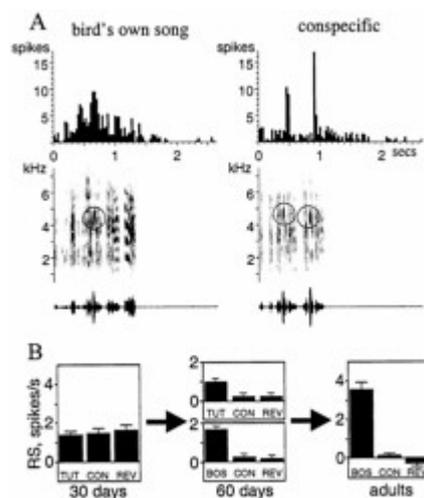


Fig. 2. AFP neurons are song-selective. (A) A song-selective neuron from an adult zebra finch. Peristimulus time histograms (PSTHs) show the greater response of a single LMAN neuron to BOS than to conspecific song. Song is shown underneath each PSTH as a sonogram (plot of frequency versus time, with the energy of each frequency band indicated by the darkness of the trace). Song-selective neurons respond to multiple acoustic features of the BOS: the circles in the sonograms identify a feature that is shared between both songs shown here and appears to elicit a response, but the figure also illustrates that many other features of BOS must contribute to the overall response of this neuron to BOS. (B) AFP neurons develop selectivity for song during development. In zebra finches of 30 days of age, LMAN neurons exhibit equivalent response strengths (RS; mean stimulus-evoked response minus background) to tutor song (TUT), conspecific song (CON), and reverse tutor song (REV). By 60 days of age, these neurons respond significantly more to TUT than to CON or to REV. In addition, BOS also elicits a much stronger response than CON and reverse BOS (REV). In adults, LMAN neurons are extremely selective for BOS.

To induce abnormal song, we bilaterally transected the tracheosyringeal portion of the hypoglossal nerve (NXII_{ts}) before song onset (\approx 25 days of age in zebra finches; Fig. 1A), thus denervating the muscles of the avian vocal organ (the syrinx). These juveniles therefore experienced a normal sensory phase with their tutor, but their entire experience of BOS was of the

abnormal, nerve cut (ts cut) song. Song analyses demonstrated that this manipulation successfully minimized both the spectral and temporal similarity between BOS and tutor song.

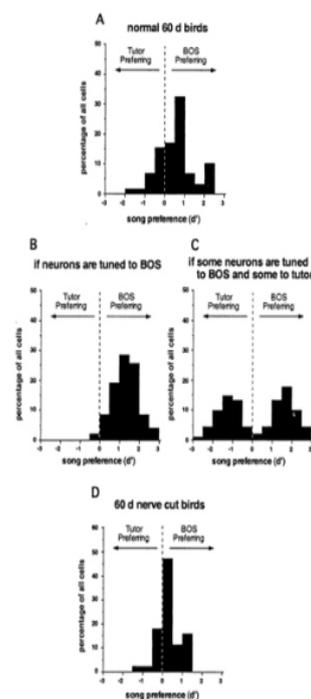


Fig. 3. Preferences for BOS versus tutor song by single AF neurons. (A) Histograms show that in 60-day-old zebra finches, there is a range of preferences among LMAN neurons. The preference for each neuron is quantified with a d' value (see refs. 18 and 25). When $d' \geq 0.5$, this indicates a strong preference for BOS over tutor song; when $d' \leq -0.5$, this indicates a strong preference for tutor song over BOS. Neurons with d' values in between were considered to have equivalent responses to both song stimuli. (B and C) Predicted results of the manipulation of BOS. (B) If neurons with equivalent responses to BOS and tutor song are shaped by BOS during development but respond to both stimuli as a result of acoustic similarities between these two songs, this type of dually responsive neuron is not expected in birds with songs unlike their tutor song, and the distribution should reveal only BOS-tuned neurons. If both BOS and tutor song independently shape different neurons in the AFP, the distribution in birds with songs very different from their tutor songs is predicted to be bimodal, as shown by the histogram. (D) The observed distribution of song preferences from ts cut birds at 60 days of age. Neurons with equivalent responses to BOS and tutor song were maintained, even though these birds' songs did not resemble the tutor song.

Using ts cut song and tutor song as stimuli, we characterized neuronal selectivity in the AFP of ts cut birds at 60 days of age. Some neurons responded more strongly to the unique ts cut BOS (tsBOS) than to tutor song, clearly demonstrating a role for BOS experience in shaping neural selectivity. Strikingly, a sizable proportion of neurons still responded equally well to both tsBOS and tutor song, despite the acoustic differences between these two songs (Fig. 3D). These neurons were not simply immature, because they exhibited selectivity for tsBOS and tutor song over conspecific and reverse song. Thus, the presence of neurons with equivalent responses to tsBOS and tutor song in these ts cut birds suggests that both song experiences can shape the selectivity of single neurons.

How might these different types of song selectivity function in song learning? Because BOS selectivity reflects the bird's current vocal output, it might provide information about the state of plastic song to a neural circuit involved in comparing BOS to a tutor song template stored in sensory coordinates. The high selectivity for BOS also might provide a kind of filter or gating function, aiding the bird in distinguishing its own vocalizations from those of others. It also may reflect in some way the pattern of motor activation during singing. The function of this selectivity could be further investigated with experiments in which AFP selectivity was broadened during song learning, perhaps with pharmacological agents.

Tutor song selectivity could encode information about the tutor song and function during sensorimotor learning as the neural reference of tutor song for birds. That is, this selectivity would result from experience of the tutor song during the sensory phase of learning. During the sensorimotor phase, the level or pattern of firing of these neurons in response to BOS then would reflect the degree to which BOS resembles the tutor song. A role for the AFP in sensory learning of the template also is supported by behavioral experiments that demonstrate a need for normal LMAN activity specifically during tutor song exposure (23).

In addition, these experiments found that BOS selectivity often coexists with tutor song selectivity in the same individual AFP neurons. This dual selectivity may reflect a function for AFP neurons in the actual comparison of BOS and tutor song that is essential to learning. For example, auditory feedback from the bird's own vocalizations would elicit activity from BOS selective cells. If this auditory feedback of the bird's own voice also matched the tutor song, then this might elicit greater or different activity in neurons that were also tuned to the tutor song than in neurons tuned to BOS or tutor song alone. Thus, the extent to which BOS resembles the tutor might be reflected in activity of dually tuned neurons, which then could participate in the reinforcement of the motor pathway.

A further suggestion that song selectivity might not only be linked to evaluation of auditory feedback, but might actually be sensitive to how well that feedback matches the target, came from studies of adult birds that were experimentally prevented from ever producing a good copy of their tutor template. We found that when we let birds that experienced transections of the tracheosyringeal portion of the hypoglossal nerve before song onset grow to adulthood, they had abnormally low song selectivity in the AFP (24). Neurons were selective enough to discriminate BOS and tutor song from conspecific and reverse songs, but the degree of selectivity was less than that found in normal adults. This result suggests that selectivity is compromised by a chronic inability of birds to match their tutor song model. If true, then these neurons are not simply reflecting sensory experience, but are influenced by the degree of matching during sensorimotor learning. Similarly, LMAN selectivity is not apparent in adult birds raised in isolation, even though they have developed stereotyped songs.[‡] Isolate birds do not have experience with a tutor, so they may experience a mismatch between

[‡]Maekawa, M. (1998) *soc. Neurosci. Abstr.* 24, 190.

auditory feedback of BOS and the template because they are missing this internal song model.

Despite the joint representation of BOS and tutor song in many AFP neurons, it seems likely that a pure sensory representation of tutor song is present somewhere in the brain. Although this could be encoded by an unidentified subset of neurons lying within other song system nuclei or even within the AFP, it seems equally plausible that such a representation lies elsewhere in the brain, perhaps in the earlier high-level auditory areas that also process songs of conspecifics (26, 27).

Song Maintenance in Adults

The striking selectivity for BOS and tutor song found in the AFP, along with the knowledge that lesions to this circuit disrupt song learning (12–14, 23), are consistent with a role for the AFP in evaluation of vocal output and adaptive modification of song during sensorimotor learning. An apparent problem for the proposed function for the AFP arose when it became clear that adult zebra finches that have been deafened or that experience persistently altered feedback of their own vocalizations exhibit gradual deterioration of their song structure (5, 10). In contrast, lesions of the AFP have no effect on normal adult song production (28). These results clearly indicate that lesions of the AFP are not equivalent to interrupting the neural encoding of auditory feedback as it enters the nervous system. They do not, however, rule out a role for the AFP in evaluating auditory feedback of song in adult birds (29).

The distinction between encoding and evaluating auditory feedback of song is illustrated in Fig. 4 A–D, which traces the hypothetical flow of auditory feedback through the song system. Normally, during song learning, auditory feedback of BOS is first encoded by the nervous system, and then compared with an internal model of the song target, the template. To the extent that there is a mismatch between the bird's song and the template, the evaluation results in an instructive signal that drives adaptive changes in the song motor program (Fig. 4A). After the completion of learning, auditory feedback of adult song matches the template, and evaluation results in a small or stable signal that does not cause changes in song (Fig. 4B). By this model, the stability of adult song results from a match between vocal output and the internal target, rather than loss of sensitivity of the motor pathway to auditory experience. Subsequent removal of all auditory feedback, however, causes a large mismatch between expected and actual feedback, and would again generate an instructive signal that drives (now nonadaptive) changes in the song (Fig. 4C), such as those seen experimentally in deaf zebra finches (5). In contrast, if instead of interrupting auditory feedback before its evaluation, the output of the evaluation itself were interrupted (Fig. 4D), the consequences would be quite different. In this case, no instructive signal would reach the motor pathway and song would not change. Indeed, by this hypothesis, interrupting the instructive signal would occlude the effects of altering auditory feedback (Fig. 4D).

In the context of this model, the failure of LMAN lesions to cause deterioration in song (28) indicates that these lesions do not interrupt auditory feedback before its evaluation. We tested the alternate hypothesis, that the AFP participates in evaluating auditory feedback, by comparing changes in song that followed deafening with those that followed a combination of deafening and bilateral AFP lesions directed at nucleus LMAN. If LMAN has a role in guiding vocal adaptation based on auditory feedback, then we expected that LMAN lesions would block the effects of deafening.

In agreement with other studies (5, 6), the songs of intact adult zebra finches were stable over long periods of time, whereas the songs of birds deafened as adults gradually deteriorated over a period of weeks to months. This deterioration included severe changes in both syllable structure and temporal patterning of song. In contrast, the songs of birds that received LMAN lesions at the same time as they were deafened remained stable and essentially unchanged (Fig. 4E). In fact, their songs were as stable as those of intact adults and remained so for at least a year, indicating that lesions did not simply delay the effects of deafening. Thus, LMAN lesions blocked deafening-induced song deterioration, including changes to both syllable and temporal structure.

These results demonstrate that the deterioration of song after deafening is an active process, because it can be blocked by lesions of a particular brain area. Moreover, these findings indicate that LMAN is required for changes in adult song, consistent with the AFP either computing or conveying an instructive signal about the quality of song, which then drives changes in vocal output. This interpretation is consistent with two previously reported effects of adult lesions of LMAN: the prevention of incorporation of new syllables into songs of birds undergoing late learning (30), and the prevention of gradual changes to the abnormal songs of adult birds whose motor production has been disrupted by denervation of the syrinx (31). In both these cases LMAN lesions also may act by eliminating signals from the AFP to the motor pathway about the mismatch between sensory feedback and the stored song model, thereby eliminating any impetus for change in vocal output. Finally, our results are compatible with the AFP playing the same role in adults as hypothesized for juveniles: the auditory feedback-based evaluation and adaptive modification of the BOS. The difference between LMAN lesions in juveniles and those in adults does not reflect a changing function of LMAN during development; rather, it highlights the difference between the state of the motor pathway in plastic song and adult song. When vocal output is in a state of flux, as during learning, the sudden absence of an instructive signal leads to a failure of song progression. In contrast, when song is already stable, as in adults, the presumptive instructive signal for change is small, and its removal has no effect. The continuing function of the AFP in adult birds is revealed only when a mismatch between actual and expected feedback is experimentally generated. Presumably, in adult birds, the AFP normally participates in the correction of any alterations in song that result from small changes to the motor pathway. Such changes to the motor pathway must normally be remarkably small, because auditory feedback is not required (in birds with AFP lesions) for adult song to remain stable.

An alternate (and not mutually exclusive) hypothesis about the role of the AFP is that neural or trophic inputs from this circuit are permissive for plasticity in the motor pathway (32, 33). According to this hypothesis, the instructive signal driving changes in song might arise elsewhere, but without the integrity of the AFP would be unable to bring about changes in song. The question of whether AFP is instructive, permissive, or both in driving vocal plasticity requires further investigation, perhaps with experiments in which the pattern, but not the amount, of AFP activity is altered.

Exactly how this putative instructive signal would be encoded by neural activity is unclear. Its sign is unknown, and it could manifest itself as magnitude and/or patterns of activity. Experiments to address these issues could assess whether AFP neural activity changes when the normal match between what the bird intends to produce and what it actually hears is disrupted. Manipulations in which auditory feedback is consistently altered in a way to which the bird gradually could adapt by altering its vocalizations would be especially informative, because changes in the instructive signal during altered feedback could be associated with subsequent changes in vocal output. For example, an initial change in auditory feedback of the song (for instance an upward shift in frequency) should result in a change in activity in neurons that compute an instructive signal, signifying a mismatch between auditory feedback and the internal song

target. After a period of this altered feedback, the bird might gradually change its vocal output (in this instance lowering the frequency of its song), so that the auditory feedback once again resembles the internal song model. During this adaptation, the activity of an instructive signal gradually should return to its original condition. Psychophysical experiments in humans show just such rapid adaptive changes in vocal behavior in response to altered feedback (34), but the song system provides the opportunity to record not only the vocal output but also the neural signals in the AFP and motor pathways during such experiments.

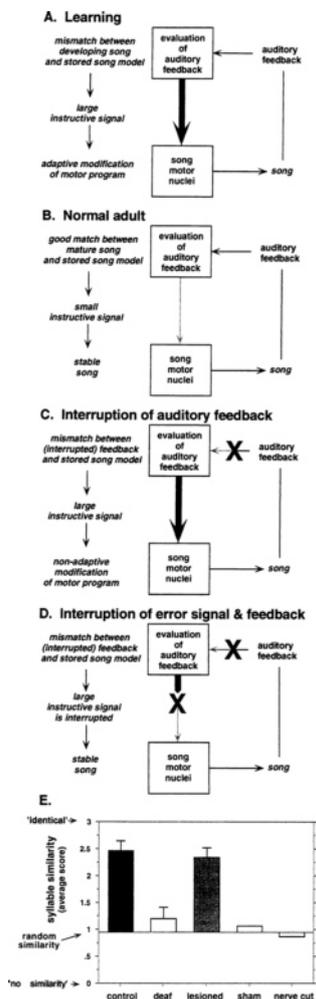


Fig. 4. Hypothetical flow of auditory feedback through the song system during sensorimotor processes and effect of AFP lesions. (A) During learning, auditory feedback from song is compared with an internal song model. Differences between the bird's own vocalizations and the model result in an instructive signal that drives adaptive modification of song. (B) In adulthood, the bird's song resembles the stored song model. Thus, there is little drive for vocal change. (C) Removal of auditory feedback leads to generation of a large instructive signal that, because of inappropriate feedback, leads to nonadaptive changes in song. (D) Interruption of the instructive signal removes the drive for change in song, even when auditory feedback is eliminated. (E) Comparison of syllable stability among adult birds that were deafened, those that were deafened and had also received LMAN lesions, and intact controls. Two additional controls included deaf birds that received sham lesions, and hearing birds with acute nerve transections, which eliminate syllable structure. The similarity of song syllables measured the degree to which syllables produced before experimental manipulations resembled those produced afterward (see ref. 29). Bars show means and standard errors for each group.

Singing-Related Activity in the AFP

Because the auditory feedback most relevant to song learning and maintenance occurs when the bird actually sings, it was clearly critical to record AFP activity during singing. To characterize signals present in the AFP of normal adult birds, we recorded single-unit and multiunit activity in LMAN during singing in adult zebra finches (35, 36). LMAN neurons fired vigorously throughout singing in adult birds (Fig. 5), despite the fact that this nucleus is not required for normal song production. Moreover, excitation began before song output, indicating that at least some of the activity is independent of auditory feedback of the bird's own voice. On average there was a consistent pattern of activity related to individual song elements, and peaks of activity tended to precede syllables. This activity resembles that reported in previous studies of singing-related premotor activity in the song control nucleus HVC (37, 38). This raises the possibility that much of the AFP activity during singing originates from the song motor circuit and may represent in part a version of the premotor signals also sent to the motor output pathway.

The properties of this singing-related activity raised the question of whether any of it is related to sensory feedback. In playback experiments, song-selective responses to auditory stimuli like those studied in anesthetized birds were apparent in LMAN of awake birds. However, they were variable from trial to trial and between birds, and it remains to be determined whether they are present in the same neurons that show singing-related activity. Moreover, the level of activity elicited by playback of auditory stimuli was low relative to singing-related activity, making it possible that small auditory feedback signals are embedded within the robust singing-related AFP activity. As an initial step to see whether AFP activity during singing contains both sensory and motor activity, we recorded multiunit activity from LMAN before and 1–3 days after deafening. Neural activity during singing was very similar predeafening and postdeafening, indicating that much of the activity during singing is not dramatically altered by an acute loss of auditory feedback.

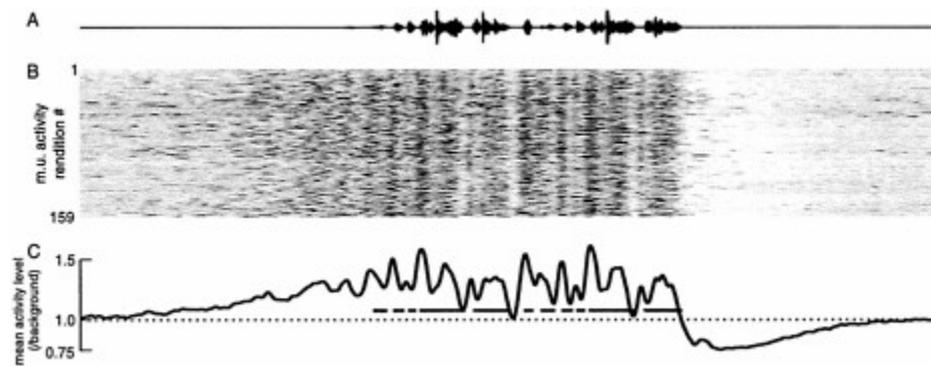


Fig. 5. LMAN neurons exhibit strong, singing-related activity. The oscillogram shows the song produced by the bird; the mean level of LMAN multiunit activity recorded in this bird before, during, and after each of 159 renditions of the song is shown below, aligned to the song. Activity level is represented by a color scale, where black indicates high neural activity, and white low activity (see ref. 36). The bottom trace shows the mean of activity during all of the renditions above, illustrating the onset of AFP activity before sound, and the peaks of activity related to syllables, which are indicated by the black bars. The duration of the entire panel is 4.5 sec.

Although selective responses to playback of BOS had suggested that sensorimotor learning influences the AFP, the marked AFP activity during singing demonstrates very directly that this circuit is not a pure sensory pathway, but instead, a sensorimotor circuit. Its function during singing may be clarified with studies that determine whether activity in individual neurons or across a population of neurons is a mixture of motor and sensory signals, and if so, how these relate to each other. In addition, recording LMAN activity in response to altered rather than absent feedback could be an important approach to studying these neurons; this would allow multiple interleaved recordings of song-related activity with and without altered feedback, which could be useful for detecting small sensory feedback signals.

The singing-related activity in the AFP might represent an efference copy, perhaps predicting the sensory consequences of motor commands. The properties of AFP neurons are consistent with this hypothesis. Because efference copy signals are triggered by motor commands, neurons with such signals would be expected to be active during singing, even in the absence of auditory feedback. Furthermore, if these neurons encode an internally generated prediction of the sensory outcome of a motor command to sing, then they might exhibit BOS selectivity when probed with song stimuli in playback experiments. Efference copies often are seen in sensorimotor systems (39–41) and can be useful for providing information about intended motor activity to multiple areas of the brain and for comparing motor instructions with the consequences of these instructions. The utility of an efference copy signal during sensorimotor learning has been explored in a computational model (42, 43). In this model, premotor activity in HVC gradually becomes associated with the resulting auditory feedback. This creates an internal prediction of the auditory feedback expected after a particular motor command is elicited. Thus, this efference copy is learned, and the role of auditory feedback is to maintain an accurate efference copy. The AFP then evaluates this sensory prediction, rather than the actual feedback. One advantage of this scheme is that it greatly shortens the normal delay between motor activity and auditory feedback, which otherwise might cause feedback evaluation signals to arrive during the motor commands for the next vocal gesture.

If a sensory prediction is learned as described above, then the considerable time it takes for altered auditory feedback to result in vocal change in adult birds might reflect the time necessary to revise the efference copy signal. An instructive signal for change would emerge only after consistently altered feedback changed the pattern of association of auditory feedback and motor commands in HVC. Alternatively, the time course for vocal change after deafening could reflect the time necessary for an instructive signal to take effect within the motor pathway. In this scheme, altered auditory feedback would immediately result in an instructive signal; however, a change in vocal output would not occur until the instructive signal was maintained over a certain period. Simultaneous recordings in the AFP and the motor pathway during singing and especially during song learning should help clarify the relationship of AFP activity to motor output and sensory feedback. Presentation of incorrect feedback again might be a useful manipulation, because altered feedback should provide a more potent signal for altering the association between motor commands and feedback in the putative efference copy than the complete absence of sound. In humans, delayed or altered auditory feedback changes vocal output much more rapidly than deafness (3, 34, 44).

Conclusions

The studies here used a combination of neurophysiological and behavioral studies to investigate the function of the AFP, in each case investigating not only normal birds but also animals in which the normal relationship between vocal motor output and sensory input had been in some way disrupted. The results revealed that AFP neurons develop selectivity during learning for both BOS and tutor song, are required for changes in vocal output throughout the bird's life, and are strongly active during singing, even in deaf birds. This suggests that AFP neurons reflect multiple sensory and motor aspects of song, and that these processes are almost inextricably entangled, even at the level of single neurons. In this respect, bird song is reminiscent of human speech (1): electrical stimulation of a single language area can affect both production and perception of speech (45), and some cortical neurons respond differently to the same word depending on whether it was spoken by the subject or by someone else (46). Perhaps this entanglement indicates that the primary task assigned to the song system, and to many speech areas as well, is

not sensory learning, but rather the sensorimotor learning required to produce a vocal imitation. Memorizing a tutor song is crucial to learning, but can be very rapid (47–49). In contrast, developing a learned song, like other motor skills, is a protracted process, involving a series of interactions between motor acts, feedback from those acts, and gradual convergence onto an internalized song model. Thus, sensorimotor learning of song alone may be sufficient to have created the need for the song system, and to have specialized it for sensorimotor processing, with much of the initial sensory processing and memorizing of songs taking place elsewhere in the brain.

It is also relevant to the results here that the AFP is a cortical-basal ganglia circuit (50, 51). Such basal ganglia circuits are well conserved evolutionarily and generally are implicated in motor and reinforcement learning, functions critical to sensorimotor learning of song. In primates, striatal neurons have predictive information related to movement and reward and might participate in comparisons of motor output to internal models (52–54). AFP neurons could similarly receive or even compute reinforcement signals and transfer them to the motor pathway. Moreover, the covert contribution of the AFP to adult plasticity is reminiscent of mammalian systems, where damage to cortical-basal ganglia circuits can impair procedural learning while having little effect on previously learned performance (55–57). Because the AFP is a discrete basal ganglia-forebrain circuit specialized for one well-defined behavior, it may prove a particularly tractable system for elucidating the neural signals present in these structures and their function in the learning and modification of sequenced motor acts.

The work described here was supported by the National Institutes of Health (Grants MH55987 and NS34835 to A.J.D. and Grant NS00913 to N.A.H.), a National Sciences Foundation graduate fellowship (M.M.S.), a Burroughs Wellcome Fund fellowship of the Life Sciences Research Foundation (M.S.B.), and the John Merck Fund and the EJLB Foundation (A.J.D.).

1. Doupe, A.J. & Kuhl, P.K. (1999) *Annu. Rev. Neurosci.* **22**, 567–631.
2. Waldstein, R.S. (1989) *J. Acoust. Soc. Am.* **88**, 2099–2114.
3. Cowie, R. & Douglas-Cowie, E. (1992) *Postlingually Acquired Deafness: Speech Deterioration and the Wider Consequences*, ed. Winter, W. (de Gruyter, Berlin).
4. Konishi, M. (1965) *Z.Tierpsychol* **22**, 770–783.
5. Nordeen, K.W. & Nordeen, E.J. (1992) *Behav. Neural Biol.* **57**, 58–66.
6. Price, P.H. (1979) *J. Comp. Physiol. Psychol.* **93**, 268–277.
7. Marler, P. (1970) *J. Comp. Physiol. Psychol.* **71**, 1–25.
8. Immelmann, K. (1969) in *Bird Vocalizations*, ed. Hinde, R.A. (Cambridge Univ. Press, London), pp. 61–74.
9. Eales, L.A. (1985) *Anim. Behav.* **33**, 1293–1300.
10. Leonardo, A. & Konishi, M. (1999) *Nature (London)* **399**, 466–470.
11. Nottebohm, F., Stokes, T.M. & Leonard, C.M. (1976) *J. Comp. Neurol.* **165**, 457–486.
12. Bottjer, S.W., Miesner, E.A. & Arnold, A.P. (1984) *Science* **224**, 901–903.
13. Sohrabji, F., Nordeen, E.J. & Nordeen, K.W. (1990) *Behav. Neural Biol.* **53**, 51–63.
14. Scharff, C. & Nottebohm, F. (1991) *J. Neurosci.* **11**, 2896–2913.
15. Doupe, A.J. (1997) *J. Neurosci.* **17**, 1147–1167.
16. McCasland, J.S. & Konishi, M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7815–7819.
17. Margoliash, D. (1983) *J. Neurosci.* **3**, 1039–1057.
18. Solis, M.M. & Doupe, A.J. (1997) *J. Neurosci.* **17**, 6447–6462.
19. Eimas, P.D., Miller, J.L. & Jusczyk, P.W. (1987) in *Categorical Perception*, ed. Hamard, S. (Cambridge Univ. Press, New York), pp. 161–195.
20. Kuhl, P.K. (1994) *Curr. Opin. Neurobiol.* **4**, 812–822.
21. Werker, J.F. & Tees, R.C. (1992) *Annu. Rev. Neurosci.* **15**, 377–402.
22. Solis, M.M. & Doupe, A.J. (1999) *J. Neurosci.* **19**, 4559–4584.
23. Basham, M.E., Nordeen, E.J. & Nordeen, K.W. (1996) *Neurobiol. Learning Mem.* **66**, 295–304.
24. Solis, M.M. & Doupe, A.J. (2000) *Neuron* **25**, 109–121.
25. Theunissen, F.E. & Doupe, A.J. (1998) *J. Neurosci.* **18**, 3786–3802.
26. Mello, C.V., Vicario, D.S. & Clayton, D.F. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6818–6822.
27. Bolhuis, J.J., Zijlstra, G.G.O., den Boer-Visser, A.M. & Van der Zee, E.A. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 2282–2285.
28. Nordeen, K.W. & Nordeen, E.J. (1993) *Behav. Neural Biol.* **59**, 79–82.
29. Brainard, M.S. & Doupe, A.J. (2000) *Nature (London)* **404**, 762–766.
30. Morrison, R.G. & Nottebohm, F. (1993) *J. Neurobiol.* **24**, 1045–1064.
31. Williams, H. & Mehta, N. (1999) *J. Neurobiol.* **39**, 14–28.
32. Johnson, F. & Bottjer, S.W. (1994) *Development (Cambridge, U.K.)* **120**, 13–24.
33. Akutagawa, E. & Konishi, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12413–12417.
34. Houde, J.F. & Jordan, M.I. (1998) *Science* **36**, 1213–1216.
35. Hessler, N.A. & Doupe, A.J. (1999) *Nat. Neurosci.* **2**, 209–211.
36. Hessler, N.A. & Doupe, A.J. (1999) *J. Neurosci.* **19**, 10461–10481.
37. McCasland, J.S. (1987) *J. Neurosci.* **7**, 23–39.
38. Yu, A.C. & Margoliash, D. (1996) *Science* **273**, 1871–1875.
39. von Holst, E. & Mittelstaedt, H. (1950) *Naturwissenschaften* **37**, 464–476.
40. Bell, C. (1989) *J. Exp. Biol.* **146**, 229–253.
41. Bridgeman, B. (1995) *Ann. Biomed. Eng.* **23**, 409–422.
42. Troyer, T. & Doupe, A.J. (2000) *J. Neurophysiol.* **84**, 1204–1223.
43. Troyer, T. & Doupe, A.J. (2000) *J. Neurophysiol.* **84**, 1224–1239.
44. Lee, B.S. (1950) *J. Acoust. Soc. Am.* **22**, 824–826.
45. Ojemann, G.A. (1991) *J. Neurosci.* **11**, 2281–2287.
46. Creutzfeldt, O., Ojemann, G. & Lettich, E. (1989) *Exp. Brain Res.* **77**, 451–489.
47. Todt, D., Hultsch, H. & Heike, D. (1979) *Z.Tierpsychol.* **51**, 23–35.
48. Peters, S., Marler, P. & Nowicki, S. (1992) *Condor* **94**, 1016–1019.
49. Tchernichovski, O., Lints, T., Mitra, P.P. & Nottebohm, F. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 12901–12904.
50. Bottjer, S.W. & Johnson, F. (1997) *J. Neurobiol.* **33**, 602–618.
51. Luo, M. & Perkel, D.J. (1999) *J. Neurosci.* **19**, 6700–6711.
52. Hikosaka, O., Sakamoto, M. & Usui, S. (1989) *J. Neurophysiol.* **61**, 814–832.
53. Hollerman, J.R., Tremblay, L. & Schultz, W. (1998) *J. Neurophysiol.* **80**, 947–963.
54. Tremblay, L., Hollerman, J.R. & Schultz, W. (1998) *J. Neurophysiol.* **80**, 964–977.
55. Graybiel, A.M., Aosaki, T., Flaherty, A.W. & Kimura, M. (1994) *Science* **265**, 1826–1831.
56. Knowlton, B.J., Mangels, J.A. & Squire, L.R. (1996) *Science* **6**, 1399–1402.
57. Nakamura, K., Sakai, K. & Hikosaka, O. (1999) *J. Neurophysiol.* **82**, 1063–1068.

ON CORTICAL CODING OF VOCAL COMMUNICATION SOUNDS IN PRIMATES

Xiaoqin Wang*

Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, 720 Rutland Avenue, Ross 424, Baltimore, MD 21205

Understanding how the brain processes vocal communication sounds is one of the most challenging problems in neuroscience. Our understanding of how the cortex accomplishes this unique task should greatly facilitate our understanding of cortical mechanisms in general. Perception of species-specific communication sounds is an important aspect of the auditory behavior of many animal species and is crucial for their social interactions, reproductive success, and survival. The principles of neural representations of these behaviorally important sounds in the cerebral cortex have direct implications for the neural mechanisms underlying human speech perception. Our progress in this area has been relatively slow, compared with our understanding of other auditory functions such as echolocation and sound localization. This article discusses previous and current studies in this field, with emphasis on nonhuman primates, and proposes a conceptual platform to further our exploration of this frontier. It is argued that the prerequisite condition for understanding cortical mechanisms underlying communication sound perception and production is an appropriate animal model. Three issues are central to this work: (i) neural encoding of statistical structure of communication sounds, (ii) the role of behavioral relevance in shaping cortical representations, and (iii) sensory-motor interactions between vocal production and perception systems.

Communication sounds are a subset of acoustic signals vocalized by a species and used in intraspecies interactions. Human speech and species-specific vocalizations of nonhuman primates are two examples of communication sounds. Vocal repertoires of many animal species also include sounds that are not communicative in nature but are essential for the behavior of a species. For example, echolocating bats emit sonar signals that are used to determine target properties of prey (e.g., distance, velocity, etc.) but are not used in social interactions between bats. Many avian species such as songbirds have rich vocal repertoires. Communication sounds of nonhuman primates are a class of acoustic signals of special interest to us. Compared with other nonhuman species, primates share the most similarities with humans in the anatomical structures of their central nervous systems, including the cerebral cortex. Neural mechanisms operating in the cortex of primates thus have direct implications for those operating in the human brain.

Although field studies provide full access to the natural behavior of primates, it is difficult to combine them with physiological studies at the single neuron level in the same animals. The challenge is to develop appropriate primate models for laboratory studies where both vocal behavior and underlying physiological structures and mechanisms can be systematically investigated. This is a prerequisite if we ever want to understand how the brain processes vocal communication sounds at the level of single neurons. Most primates have a well-developed and sophisticated vocal repertoire in their natural habitats. However, for many larger primate species like macaque monkeys, their vocal activities diminish under the captive conditions commonly found in research institutions, in part because of the lack of a behaviorally suitable housing environment. Fortunately, some primate species such as New World monkeys (e.g., marmosets, squirrel monkeys) remain highly vocal in properly configured captive conditions. Fig. 1A shows a vocal exchange between a male and a female marmoset recorded from a captive marmoset colony. These primate species may serve as excellent models for us to study in detail their vocal behavior as well as underlying neural mechanisms in the brain.

The cerebral cortex is known to play an important role in processing species-specific vocalizations. Studies have shown that lesions of auditory cortex cause a deficit in speech comprehension in humans and discrimination of vocalizations in primates (1, 2). Anatomically, humans and primates have similar cytoarchitecture in the superior temporal gyrus where the sensory auditory cortex is located (3). It has been shown in both Old World and New World primates that the sensory auditory cortex consists of a primary auditory field (A1) and surrounding secondary fields (4–6). Afferent information on acoustic signals is processed first in A1 and then in the secondary fields. From there, the information flows to the frontal cortex and to other parts of the cerebral cortex. The issues discussed in this article concern coding of vocalizations in the superior temporal gyrus, the sensory auditory cortex.

Issues in Understanding Cortical Processing of Communication Sounds

Vocal communication sounds are a special class of signals that are characterized by their acoustical complexity, their biological importance, and the fact that they are produced and perceived by the same species. As a result, the representations of these signals in the brain are likely to be different from representations of other types of acoustic signals (e.g., behaviorally irrelevant sounds, sounds from prey and predators). The nature of vocal communication sounds requires that we consider the following issues in our studies.

Neural Encoding of Statistical Structures of Communication Sounds. There are at least three important behavioral tasks for which primates rely on their vocalizations in a natural environment. The tasks are to (i) identify messages conveyed by members of a social group or family, (ii) identify the caller of a vocalization, and (iii) determine the spatial location of a caller. There has been ample evidence that primates use their vocalizations in these behavioral tasks (7). These tasks require the auditory system of primates to solve “what”, “who,” and “where” problems on the basis of vocalizations. Before one can understand how the

*E-mail: xwang@bme.jhu.edu.

This paper was presented at the National Academy of Sciences colloquium “Auditory Neuroscience: Development, Transduction, and Integration,” held May 19–21, 2000, at the Arnold and Mabel Beckman Center in Irvine, CA.

Abbreviation: CF, characteristic frequency.

auditory system solves these problems, however, one must understand how the information is coded in the acoustic structures of vocalizations. After all, in the absence of visual and other kinds of sensory information (a realistic situation encountered by monkeys living in jungles with heavy vegetation), vocalization is the only messenger. Because spatial location of a sound is computed by the auditory system by using binaural cues as well as spectral cues provided by the external ear, the fundamental information carried by vocalizations is for “what” and “who” problems. Primates are known to produce distinct call types that presumably encode “what” information. It is also known that each animal has its own idiosyncratic features in its vocalizations, which are likely cues to represent caller identity.

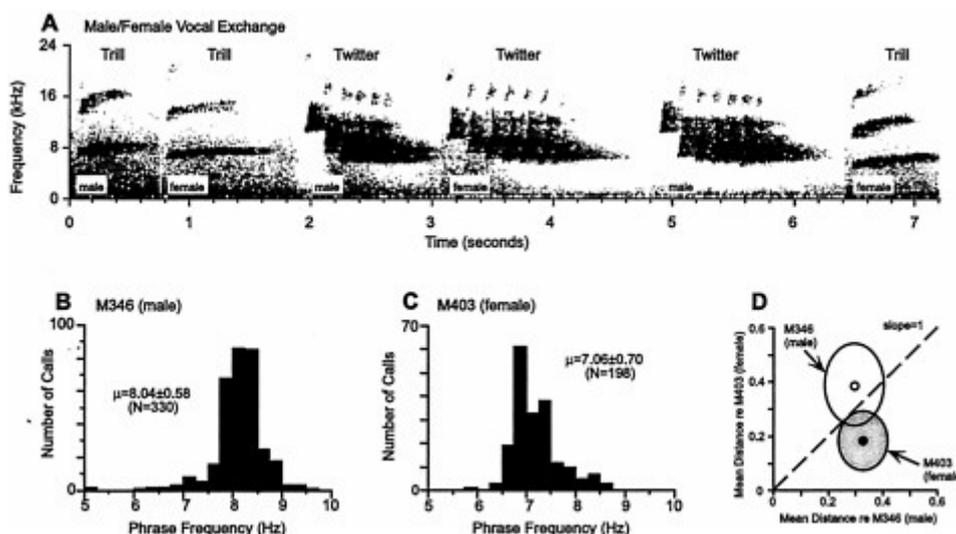


Fig. 1. (A) A real-time recording of vocal exchanges between a pair of marmosets (one male, one female), shown in the form of a spectrogram. This typical vocal exchange contains trill and twitter calls. Call type and caller identity of each call are indicated. (B and C) Distribution of *phrase frequency* of twitter calls from a male (B, M346) and female (C, M403) marmoset. Both marmosets lived in the same colony and engaged in frequent vocal exchanges such as the example shown in A. A twitter call contains several discrete upward FM sweeps, each of which is referred to as a phrase. The intervals between phrases are relatively constant in each twitter call. Fourier analysis of a twitter call’s envelope, obtained by using the Hilbert transform, revealed a local maximum reflecting the repetition frequency of the phrases. The frequency at this maximum is defined as the phrase frequency (27). (D) The results of a multidimensional clustering analysis of twitter calls of the pair of marmoset monkeys in B and C. Four parameters were calculated for each twitter call sample and used in the analysis. These parameters included (i) number of phrases, (ii) phrase frequency, (iii) spectral-peak frequency of the first phrase, and (iv) spectral-peak frequency of the second phrase. The spectral-peak frequency is computed from the magnitude spectrum of each phrase (27). Two ellipsoids are drawn by using the mean and standard deviation of a distance measure d_{ij} (see definition below). The open circle marks the mean distances of calls made by the male monkey (M346) with respect to its own group mean (abscissa) and the female monkey’s group mean (ordinate); the open ellipsoid (male monkey, M346) outlines the standard deviations along both axes. The filled circle and shaded ellipsoid are calculations for calls from the female monkey M403.

where i, j are animal designation (1:M346, 2:M403); N_i is number of call samples from i th animal ($N_1=330$, $N_2=198$); $P_{ik}(n)$ is k th parameter in the n th call of the i th animal; and \bar{m}_{ik} is mean value of the k th parameter of call samples from i th animal.

However, just like human speech sounds, an individual utterance by a primate does not fully represent the underlying structure of a particular call type or caller because of the jittering of the vocal production apparatus and noises in the environment. There is some degree of randomness reflected by the nature of all sound production systems. The underlying statistical properties of these vocalizations, however, ought to be invariant from call to call. Thus, the corresponding cortical coding should be and can only be fully understood at this level. In other words, it is not enough to know how cortical neurons respond to a particular call sample. One must know how the cortex extracts the invariant statistical structure from which all utterances of a call type or caller are generated.

The Role of Behavioral Relevance in Shaping Cortical Representations. As biologically important signals, species-specific vocalizations have an ensured representation in the brain. To understand how these types of sound are represented in the cortex and how one can generalize principles from these representational schemes, one needs to understand the biological basis that shapes the cortical responses. There are three factors that may bias cortical responses to communication sounds in a particular species. These factors are (i) evolutionary predisposition, (ii) developmental plasticity, and (iii) experience-dependent plasticity in adulthood, each of which differs in its time scale. In a primate species whose vocal spectrum occupies much higher frequency range than that of humans, its auditory cortex devotes a larger

portion to that frequency range than does human auditory cortex. For example, human speech is concentrated on 0.5–3 kHz (the range of formant frequency), whereas vocalizations of marmoset monkeys are centered at 7–15 kHz (8). As a result, marmosets have an expanded representation of 7–15 kHz in their auditory cortex, whereas representation of frequencies below 3 kHz is much more limited (9). Such species-specific differences are formed through evolution over many millions of years and likely have a genetic underpinning. On a shorter time scale, changes throughout the developmental period, both prenatal and postnatal, may influence how the cortex processes vocalizations, given what is known of developmental plasticity of the cerebral cortex from studies of other sensory cortices. Finally, because the cortex is known to be subject to experience-dependent plasticity in adulthood (10, 11), what an animal hears on a daily basis must shape its cortical representation. The time scale for such changes is the shortest, probably in terms of months, weeks, or even days.

What do these considerations tell us in studying vocal communication sounds? For one, it is clear that one has to study directly mechanisms for encoding vocalizations within a species. However, because of developmental and experience-dependent plasticity, one may not fully reveal cortical coding mechanisms when using calls from unrelated conspecifics to which the experimental subjects have never been exposed. A simple but useful analogy is that one cannot study how the cortex codes Chinese in a native English speaker who never learned Chinese. Nor can one study cortical coding of Chinese in one of Chinese descent who was never exposed to Chinese. Although it is not yet clear to what extent these analogies are true in primates, they are powerful reminders that one must pay close attention to the behavioral meaning of sounds in any model systems of primates.

Sensory-Motor Interactions Between Vocal Production and Perception Systems. One important distinction between vocal communication sounds and other biologically important but nonvocal signals (such as sounds by prey and predators) is that the former are produced by the species perceiving them. It has long been known in studies of human speech that our perception of speech is biased by the nature of our ability to produce speech (12). Recent imaging studies showed that cortical areas outside the sensory auditory cortex on the superior temporal gyrus are activated during passive listening experience (13, 14). These findings are consistent with observations from studies in human epileptics undergoing neurological treatment with subdural or depth electrodes, through which electrical stimulation of the frontal cortex interrupted a patient's ability in listening comprehension tasks (1, 15). This evidence suggests that processing of vocal communication sounds involves both sensory and motor systems. The frontal cortex likely plays an important role in these sensory-vocal interactions. It has been shown that anatomically the frontal cortex is reciprocally connected with auditory sensory cortex on the superior temporal gyrus in both Old World and New World primates (5, 16, 17). Specifically, the connections originate from the secondary auditory fields where strong neural responses to vocalizations were found in primates [ref. 18; X.W. (1999) *Association of Research in Otolaryngology Abs.* 22, 173]. It is likely that feedback from these higher-order processing centers outside the superior temporal gyrus influences cortical coding in the sensory auditory cortex.

Other important issues in cortical coding include whether vocalizations are represented by neural ensembles or by specialized cells and how representation of vocalizations is transformed from the primary auditory cortex to the secondary areas, which will be discussed below.

Progress and Current Work

Previous Studies. Over the past several decades, a number of experimental attempts have been made to elucidate the mechanisms of cortical coding of species-specific vocalizations in the auditory cortex of primates. The effort to understand cortical representation of vocalizations reached a peak in the 1970s, when a number of reports were published. The results of these studies were mixed, with no clear or consistent picture emerging as to how communication sounds are represented in the auditory cortex of primates (19). This lack of success may be accounted for in part, retrospectively, by expectations of the form of cortical coding of behaviorally important stimuli. For a time, it was thought that primate vocalizations were encoded by highly specialized neurons, the so-called "call detectors" (20, 21). However, individual neurons in the auditory cortex were often found to respond to more than one call or to various features of calls (20–23). The initial estimate of the percentage of the "call detectors" was relatively high. Later, much smaller numbers were reported as more calls were tested and more neurons were studied. At the end of this series of explorations, it appeared, at least in the initial stages of auditory cortical pathway, that the notion of highly specialized neurons is doubtful. Perhaps because of these seemingly disappointing findings, no systematic studies on this subject were reported for more than a decade afterward.

One shortcoming in the earlier studies was that responses to vocalizations were not adequately related to basic functional properties of a neuron such as its receptive field, temporal dynamics, etc. In addition, responses to vocalizations were not interpreted in the context of the overall organization of a cortical field or overall structure of the entire auditory cortex. Such information became available only in later years, and much of the structure of the auditory cortex is still being defined in primates. It is clear now that until a good understanding of the properties of a neuron is achieved, besides its responses to vocalizations, cortical coding of species-specific vocalizations will not be fully understood. Another lesson learned from the earlier studies is that one must consider the underlying statistical structure of a species' vocalizations. In earlier studies, only vocalization tokens were used. This made it difficult, if not impossible, to accurately interpret cortical responses, as argued earlier in this article. One has to notice that in those earlier days, before powerful computers and digital technology became available, quantifying primate vocalizations would have been a formidable task. Nonetheless, these earlier explorations of the primate auditory cortex by using species-specific vocalization served as an important stepping stone for current and future studies of this important problem in auditory neurophysiology; they pointed to the right direction for seeking the correct answers.

Quantitative Characterizations of Communication Sound Repertoire. After a long period of silence, studies in this field became active again in recent years. On the forefront of understanding acoustic structure of communication sounds, my laboratory has systematically studied the vocal repertoire of communication sounds in a highly vocal primate species, the common marmoset (*Callithrix jacchus jacchus*). In this study, we quantitatively characterized statistical properties not only of various call types but also of acoustic features related to individual identity. Fig. 1 B-D illustrates the ideas of this analysis, which was based on extensive samples of vocalizations from a large colony of marmosets at Johns Hopkins University (Baltimore, MD) [ref. 24; J.A. Agamaite & X.W. (1997) *Association of Research in Otolaryngology Abstr.* 20, 144]. The results of this study showed that (i) marmosets have discrete call types with distinct acoustic structures, and (ii) idiosyncratic vocal features of individual monkeys are quantifiable on the basis of their acoustics and are clustered in a multidimensional space (Fig. 1D). Acoustic structures of

marmoset vocalizations were found to contain sufficient information for the discrimination of both caller and gender. In earlier studies of squirrel monkeys, individual call types were identified, but the statistical structure of call types was not fully evaluated, nor were vocal features related to an individual monkey (25). Our studies in marmosets filled this gap and provided a solid basis to further the exploration of cortical coding of communication sounds. Among the published reports, the most comprehensive analysis of the vocal repertoire of communication sounds in a mammalian species was recently conducted in mustached bats in the laboratory of Nobuo Suga (26).

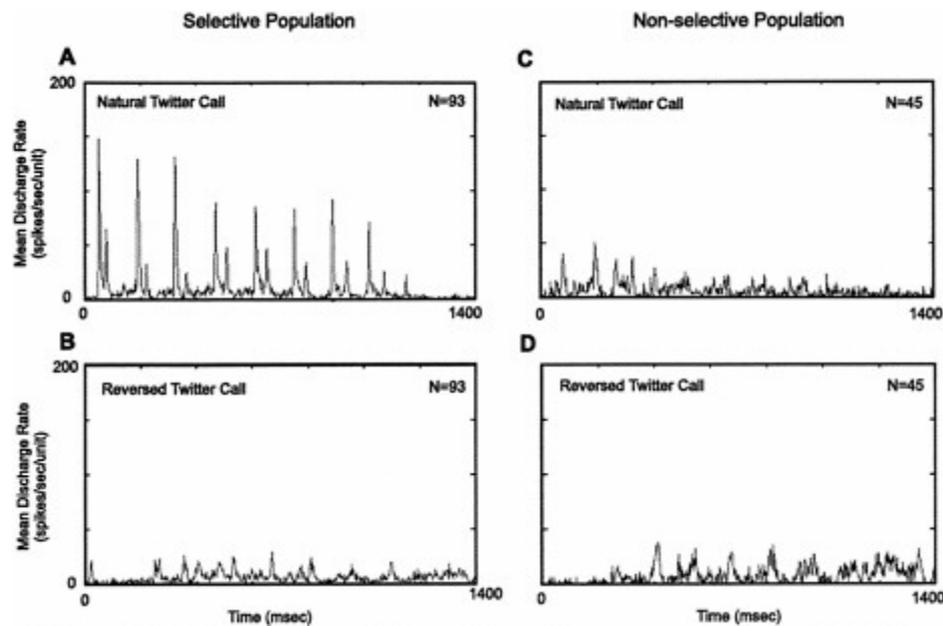


Fig. 2. Averaged temporal discharge patterns of responses to a natural twitter call and its time-reversed version recorded from the primary auditory cortex of a marmoset (27). In each of 138 sampled units, its discharge rate to the natural twitter call was compared with that to the reversed twitter call. The sampled units were divided into two subpopulation based on this analysis. Units included in the *selective population* (A, B) responded more strongly to the natural twitter call than to the reversed twitter call, whereas the units included in the *nonselective population* (C, D) responded more strongly to the reversed twitter call than to the natural twitter call. In A-D, a mean poststimulus histogram (PSTH) is shown for each neuronal population under one of the two stimulus conditions (bin width=2.0 ms).

One open issue is how the vocal repertoire recorded in captivity differs from what exists in the wild. It is likely that certain types of calls that are observed in a more behaviorally enriched natural environment are not observable in captivity. However, in the case of common marmoset, there has been no evidence that the basic acoustic structures of calls from captive animals differ fundamentally from those of corresponding calls sampled from animals in the wild. Ideally, quantitative analysis of vocalizations similar to what we have conducted in captive marmosets should be performed in future studies in a wild population of marmosets and other primate species.

Dependence of Cortical Responses on Behavioral Relevance of Vocalizations. In a recent study, it was demonstrated that natural vocalizations of marmoset monkeys produce stronger responses in the primary auditory cortex than do an equally complex but artificial sound such as a time-reversed call (27), as illustrated in Fig. 2. Moreover, the subpopulation of cortical neurons that were selective to a natural call had a more clear representation of the spectral shape of the call than did the nonselective subpopulation (Fig. 3 A and B). When the reversed call was played, responses from two populations were similar (Fig. 2 B and D and Fig. 3C). These observations suggest that marmoset auditory cortex preferentially responds to vocalizations that are of behavioral significance as compared with behaviorally irrelevant sounds.

To further test this notion, we have directly compared responses to natural and time-reversed calls in the auditory cortex of another mammal, the cat, whose A1 shares similar basic physiological properties (e.g., characteristic frequency, threshold, latency, etc.) to that of the marmoset [ref. 28; X.W., R. Beitel, C.E.Schreiner & M.M.Merzenich (1995) *Soc. Neurosci. Abstr.* **21**, 669]. Unlike marmoset, however, cat's auditory cortex does not differentiate natural marmoset vocalizations from their time-reversed version (X.W. & S.Kadia, unpublished observations). One obvious conclusion from both of these lines of evidence is that it is essential to study cortical representation of communication sounds in an appropriate model animal with relevant stimuli. This is not necessarily a trivial issue, because much of our knowledge of subcortical auditory system has been based on studies using artificial stimuli. One likely factor that contributes to the observed response differences is experience-dependent cortical plasticity. As is long known, a behaviorally important stimulus can induce profound changes in neuronal

responses of the sensory cortex even within a short period (29, 30). Such contributions would not be observed if vocalizations from unrelated conspecifics were used in obtaining cortical responses. It is possible that the difference between responses from cat and marmoset A1 are due to behavioral modulation on a larger time scale, such as differential experiences through development or even evolution. Further studies are needed to address these issues.

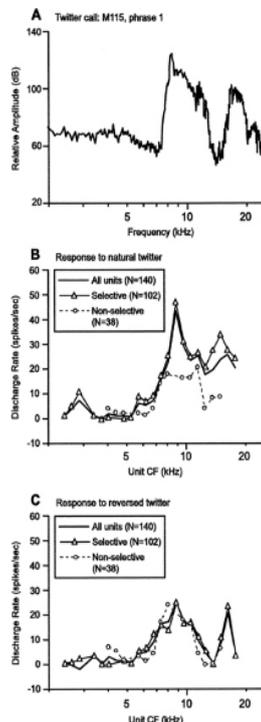


Fig. 3. Population representation of the spectral shape of marmoset vocalizations. Comparison is made between short-term call spectrum of one phrase of the twitter call and rate-CF (discharge rate vs. characteristic frequency) profiles computed over a corresponding time period. Data shown were obtained from the primary auditory cortex of one marmoset (27). (A) Magnitude spectrum of the first phrase of a natural twitter call. The magnitude spectrum of this call phrase in the time-reversed call is the same; only the phase spectrum is different. (B) Rate-CF profiles were constructed based on responses to a natural twitter call from 140 sampled units and were computed by using a triangular weighting window whose base was 0.25 octave wide. The centers of adjacent windows were 0.125 octave apart. Only averages that had at least 3 units in the window were included. Three profiles are shown, all units ($n=140$, black solid line), selective subpopulation ($n=102$, red solid line with triangle), and nonselective subpopulation ($n=38$, green dashed line with circle). The definitions of the two subpopulation of units are given in Fig. 2. (C) Rate-CF profiles are shown for cortical responses to the same call phrase as analyzed in B but delivered in the time-reversed call. The same analytic method and display format are used as in B.

The Nature of Cortical Representation. A fundamental question that has yet to be satisfactorily answered is the nature of cortical representation of species-specific vocalizations. The lack of convincing evidence of call detectors from earlier studies led us to consider other hypotheses regarding cortical coding of complex sounds such as vocalizations. An alternative strategy of encoding complex vocalizations is by the discharge patterns of spatially distributed neuronal populations. Such coding strategies have been demonstrated in the auditory nerve (31, 32) and the cochlear nucleus (33). Recent work (27) in the marmoset has provided evidence that a population coding scheme may also operate at the input stage of the auditory cortex, but in a very different form from that observed at the periphery as illustrated in Fig. 4. Fig. 4A shows population responses to three types of marmoset vocalizations. The difference between cortical and peripheral representations lies largely in the temporal characteristics of neuronal discharges. Compared with the auditory nerve, cortical neurons do not faithfully follow rapidly changing stimulus components (Fig. 4D). We recently examined this issue in the auditory cortex using click train stimuli and found that stimulus-following capacity of cortical neurons is limited to about 20–30 msec (34). Are rapid stimulus components really lost in cortical representation? The answer is no. It turned out that a subpopulation of cortical neurons responded to short interstimulus intervals with changing discharge rate (34). These neurons can potentially signal rapidly changing stimulus components.

It should be pointed out that the notion of population coding does not necessarily exclude the operation of highly specialized neurons. The questions are to what extent such specialized neurons exist and what are their functional roles? In earlier studies, it was expected that a large percentage of these neurons would be found and that the coding of species-specific vocalizations is essentially carried out by such neurons. Various studies since 1970s have cast doubts on this line of thinking. However, a recent study in mustached bats found neurons selective to the bat's communication calls in the FM-FM area of the auditory cortex (35). In the marmoset, a subpopulation of A1 neurons appeared to be more selective to alterations of natural calls than was the rest of the population (27). In our most recent studies in awake marmosets, we identified neurons that were selective to specific call types or even individual callers in both A1 and secondary cortical fields [X.W. (1999) *Association of Research in Otolaryngology Abs.* 22, 173]. Thus, there appear to be two general classes of cortical neurons, one responding selectively to call types or even callers, the other responding to a wider range of sounds including both vocalization and nonvocalization signals. The latter group accounts for a larger proportion of cortical neurons in A1 than in the lateral field (a

secondary cortical area), as observed in our experiments in awake marmosets [X.W. (1999) *Association of Research in Otolaryngology Abs.* 22, 173]. One may speculate that the task of encoding communication sounds is undertaken by the call-selective neurons alone, however small that population is. On the other hand, one cannot avoid considering the role of the nonselective population that also responds to communication sounds concurrently with the selective population. One hypothesis is that the nonselective population serves as a general sound analyzer that provides the brain with a detailed description of a sound, whereas the selective population serves as a signal classifier whose task is to tell the brain that a particular call type or caller was heard. One can imagine that the former analysis is needed when a new sound is heard or being learned, whereas the latter analysis is used when a familiar sound such as a vocalization is encountered. This and other theories (36, 37) are likely to be tested by new experiments in coming years. Nonetheless, two important lessons have been learned from earlier and recent studies. First, it is not sufficient to study individual call tokens; one must study the representation of statistical structures of communication sounds. Second, cortical representations of complex sounds like vocalizations are no longer isomorphic replicas of the acoustic spectrotemporal pattern of the sound; they have been significantly transformed into a new representational space.

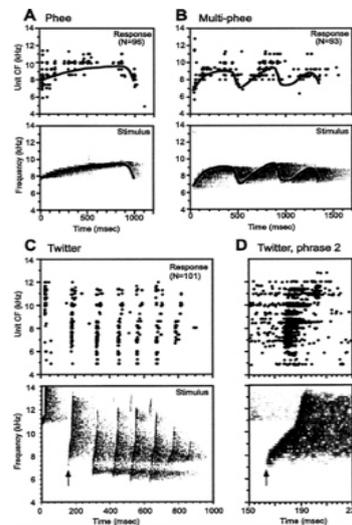


Fig. 4. Comparison between the spectrotemporal acoustic pattern of marmoset vocalizations and the corresponding spectrotemporal discharge patterns recorded in the primary auditory cortex of marmosets. In each plot (A-D), *Upper* shows population responses to a vocalization, and *Lower* shows the corresponding spectrogram of the stimulus. Discharges as they occurred in time (abscissa) from individual cortical units are aligned along the ordinate according to their objectively defined CF. The display of discharges was based on PSTHs computed for each unit (bin width=2.0 ms). All three vocalizations were delivered at the sound level of 60 dB SPL during the experiments. (A) Population responses to a marmoset *phee* call. An outline of the trajectory (solid line in red) of the call's time-varying spectral peak is drawn on *Upper* and *Lower* for comparison. (B) Population responses to a marmoset *multiple-phee* call. An outline of the trajectory (solid line in red) of the call's time-varying spectral peak is drawn on *Upper* and *Lower* for comparison. (C) Population responses to a marmoset *twitter* call. The second call phrase is indicated by a vertical arrowhead (red). (D) An expanded view of cortical responses to the second phrase of the twitter call shown in C (*Upper*) and the corresponding spectrogram of the second call phrase (*Lower*), with a time mark indicated by an arrowhead (red) as in C. Responses of the same group of cortical units shown in C are included but displayed in the form of dot raster. In *Upper*, each recorded spike occurrence within the time period shown is marked as a dot. Spike times from 10 repetitions in each unit are aligned along 10 lines centered at the CF of the unit, shifted by 10 Hz for each repetition for display purpose (i.e., positioned from CF -50 Hz to CF +40 Hz, in 10-Hz step).

A “Vocal” Pathway. A further question on cortical coding is how the information in communication sounds is routed through cortical systems, including those outside the superior temporal gyrus. It has been suggested that there are “where” and “what” pathways in the auditory cortex (6, 17), a notion largely borrowed from the visual system. However, in our opinion, the auditory cortical system in humans and primates must include another pathway, a “vocal” pathway, in processing vocal communication sounds. This pathway may or may not be involved when nonvocal acoustic signals are processed by the cerebral cortex. There is an important distinction between the auditory and visual system, however. Vocal species such as humans and primates produce their own behaviorally important inputs (i.e., speech and vocalizations) to their auditory system. As has been argued in this article, the “vocal” pathway is likely to be a mutual communication channel between the superior temporal gyrus and the frontal cortex. Anatomically, it has been shown by a number of studies in both humans and primates that the superior temporal gyrus is connected with the frontal lobe, reciprocally, where generation and control of vocal activities take place in humans and possibly in primates as well.

Future Directions. A number of issues remain to be resolved in the overall understanding of cortical coding of communication sounds. First, a careful consideration of the state of an animal must be made when its cortical responses are studied. Much of the understanding of the auditory cortex has been based on studies in anesthetized animals, which obviously carry severe limitations. An important step in moving this field forward is the use of awake and behaving preparations, which have been widely adapted in studies of the visual system. Only then can important issues like attentional modulation of cortical responses be adequately evaluated. Second, two crucial issues that must be answered in this field are vocal production mechanisms and vocal development and learning in primates. The essential question is whether primates possess voluntarily controlled and learned vocalizations. Earlier studies have painted negative pictures of both prospects. In the coming years, these earlier conclusions are likely to be challenged and possibly modified. Finally, to study these and other emerging questions on cortical coding of vocal communication sounds, new techniques are needed. The main limitation of existing neurophysiological methods is that vocal behavior of animals is substantially restricted or eliminated once an animal is restrained. Using implanted electrodes can loosen these restrictions. Such techniques are widely used in the study of rodents with tethered cables to relay signals from the electrodes to a data recorder. However, for highly mobile primates, especially when one is interested in their natural vocal activity, the

ideal way to relay signals from a recording electrode array is by means of a telemetry device.

I thank members of the Laboratory of Auditory Neurophysiology at the Biomedical Engineering Department of the Johns Hopkins University for their contributions toward the research and opinions discussed in this article, in particular James Agamaite, Thomas Lu, Siddhartha Kadia, Ross Snider, Dennis Barbour, Li Liang, Steve Eliades, and Haiyin Chen. I thank Airi Krause and Ashley Pistorio for assistance in animal training and marmoset colony management. I am especially grateful for the encouragement and support of Dr. Michael Merzenich at the Coleman Laboratory of the University of California at San Francisco, where the work on the marmoset model was initiated. I thank D.Barbour for his excellent help in graphics work and T.Lu and D.Barbour for proofreading the manuscript. This research was supported by National Institutes of Health Grant DC03180 and by a Presidential Early Career Award for Scientists and Engineers (1999–2004).

1. Penfield, W. & Roberts, L. (1959) *Speech and Brain-Mechanisms* (Princeton Univ. Press, Princeton, NJ).
2. Heffner, H.E. & Heffner, R.S. (1986) *J. Neurophysiol.* **56**, 683–701.
3. Brodmann, K. (1909) *Vergleichende Lokalisationslehre der Grobhirnrinde* (Barth, Leipzig).
4. Jones, E.G. & Powell, T.P.S. (1970) *Brain* **93**, 793–820.
5. Pandya, D.N. & Yeterian, E.H. (1985) in *Cerebral Cortex*, eds. Peters, A. & Jones, E.G. (Plenum, New York), Vol. 4, pp. 3–61.
6. Kaas, J.H., Hackett, T.A. & Tramo, M.J. (1999) *Curr. Opin. Neurobiol.* **9**, 164–170.
7. Snowdon, C.T., Brown, C.H. & Petersen, M.R. (1982) *Primate Communication* (Cambridge Univ. Press, Cambridge, U.K.).
8. Epple, G. (1968) *Folia Primatol.* **8**, 1–40.
9. Aitkin, L.M., Merzenich, M.M., Irvine, D.R., Clarey, J.C. & Nelson, J.E. (1986) *J. Comp. Neurol.* **252**, 175–185.
10. Merzenich, M.M., Nelson, R.J., Stryker, M.P., Cynader, M.S., Schoppmann, A. & Zook, J.M. (1984) *J. Comp. Neurol.* **224**, 591–605.
11. Merzenich, M.M., Kaas, J.H., Wall, J.T., Sur, M., Nelson, R.J. & Felleman, D.J. (1983) *Neuroscience* **10**, 639–665.
12. Liberman, A.M. (1996) *Speech: A Special Code* (MIT Press, Cambridge, MA).
13. Zatorre, R.J., Evans, A.C., Meyer, E. & Gjedde, A. (1992) *Science* **256**, 846–849.
14. Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M. & Prieto, T. (1997) *J. Neurosci.* **17**, 353–362.
15. Ojemann, G.A. (1991) *J. Neurosci.* **11**, 2281–2287.
16. Morel, A. & Kaas, J.H. (1992) *J. Comp. Neurol.* **318**, 27–63.
17. Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S. & Rauschecker, J.P. (1999) *Nat. Neurosci.* **2**, 1131–1136.
18. Rauschecker, J.P., Tian, B. & Hauser, M. (1995) *Science* **268**, 111–114.
19. Pelleg-Toiba, R. & Wollberg, Z. (1991) *J. Basic Clin. Physiol. Pharmacol.* **2**, 257–272.
20. Winter, P. & Funkenstein, H.H. (1973) *Exp. Brain Res.* **18**, 489–504.
21. Newman, J.D. & Wollberg, Z. (1973a) *Exp. Neurol.* **40**, 821–824.
22. Wollberg, Z. & Newman, J.D. (1972) *Science* **175**, 212–214.
23. Manley, J.A. & Muller-Preuss, P. (1978) *Exp. Brain Res.* **32**, 171–180.
24. Agamaite, J.A. (1997) M.S.E. thesis (Johns Hopkins University, Baltimore, MD).
25. Winter, P., Ploog, D. & Latta, J. (1966) *Exp Brain Res.* **1**, 359–384.
26. Kanwal, J.S., Matsumura, S., Ohlemiller, K. & Suga, N. (1994) *J. Acoust. Soc. Am.* **96**, 1229–1254.
27. Wang, X., Merzenich, M.M., Beitel, R. & Schreiner, C.E. (1995) *J. Neurophysiol.* **74**, 2685–2706.
28. Schreiner, C.E., Read, H.L. & Sutter, M.L. (2000) *Annu. Rev. Neurosci.* **23**, 501–529.
29. Recanzone, G.H., Schreiner, C.E. & Merzenich, M.M. (1992) *J. Neurosci.* **13**, 87–104.
30. Wang, X., Merzenich, M.M., Sameshima, K. & Jenkins, W.M. (1995) *Nature (London)* **378**, 71–75.
31. Sachs, M.B. & Young, E.D. (1979) *J. Acoust. Soc. Am.* **66**, 470–479.
32. Young, E.D. & Sachs, M.B. (1979) *J. Acoust. Soc. Am.* **66**, 1381–1403.
33. Blackburn, C.C. & Sachs, M.B. (1990) *J. Neurophysiol.* **63**, 1191–1212.
34. Lu, T. & Wang, X. (2000) *J. Neurophysiol.* **84**, 236–246.
35. Esser, K.H., Condon, C.J., Suga, N. & Kanwal, J.S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14019–24.
36. Suga, N. (1994) *J. Comp. Physiol. A* **175**, 135–144.
37. Suga, N. (1994) in *The Cognitive Neuroscience*, ed. Gazzanica, M.S. (MIT Press, Cambridge, MA.), pp. 295–313.

A NEW VIEW OF LANGUAGE ACQUISITION

Patricia K. Kuhl*

Department of Speech and Hearing Sciences and Center for Mind, Brain, and Learning, University of Washington, Box 357920, Seattle, WA 98195

At the forefront of debates on language are new data demonstrating infants' early acquisition of information about their native language. The data show that infants perceptually "map" critical aspects of ambient language in the first year of life before they can speak. Statistical properties of speech are picked up through exposure to ambient language. Moreover, linguistic experience alters infants' perception of speech, warping perception in the service of language. Infants' strategies are unexpected and unpredicted by historical views. A new theoretical position has emerged, and six postulates of this position are described.

The last half of the 20th century has produced a revolution in our understanding of language and its acquisition. Studies of infants across languages and cultures have provided valuable information about the initial state of the mechanisms underlying language, and more recently, have revealed infants' unexpected learning strategies. The learning strategies—demonstrating pattern perception, as well as statistical (probabilistic and distributional) computational skills—are not predicted by historical theories. The results lead to a new view of language acquisition, one that accounts for both the initial state of linguistic knowledge in infants and infants' extraordinary ability to learn simply by listening to ambient language. The new view reinterprets the critical period for language and helps explain certain paradoxes—why infants, for example, with their immature cognitive systems, far surpass adults in acquiring a new language. The goal of this paper is to illustrate the recent work and offer six principles that shape the new perspective.

Historical Theoretical Positions

In the last half of the 20th century, debate on the origins of language was ignited by a highly publicized exchange between a strong nativist and a strong learning theorist. In 1957, the behavioral psychologist B.F. Skinner proposed a learning view in his book *Verbal Behavior*, arguing that language, like all animal behavior, was an "operant" that developed in children as a function of external reinforcement and shaping (1). By Skinner's account, infants learn language as a rat learns to press a bar—through the monitoring and management of reward contingencies.

Noam Chomsky, in a review of *Verbal Behavior*, took a very different theoretical position (2, 3). Chomsky argued that traditional reinforcement learning had little to do with humans' abilities to acquire language. He posited a "language faculty" that included innately specified constraints on the possible forms human language could take. Chomsky argued that infants' innate constraints for language included specification of a universal grammar and universal phonetics. Language was one of the primary examples of what Fodor called a module—domain-specific, informationally encapsulated, and innate (4).

The two approaches took strikingly different positions on all of the critical components of a theory of language acquisition: (i) the initial state of knowledge, (ii) the mechanisms responsible for developmental change, and (iii) the role played by ambient language input. On Skinner's view, no innate information was necessary, developmental change was brought about through reward contingencies, and language input did not cause language to emerge. On Chomsky's view, infants' innate knowledge of language was a core tenet, development constituted "growth" or maturation of the language module, and language input triggered (or set the parameters for) a particular pattern from among those innately provided.

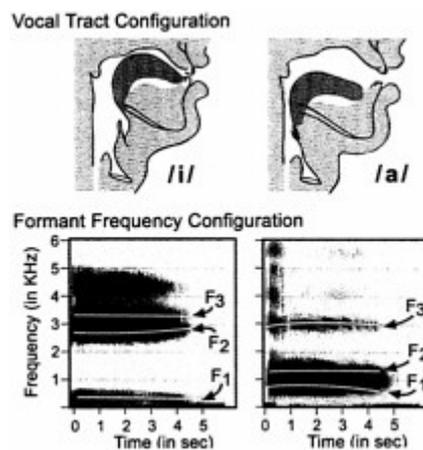


Fig. 1. Vocal tract positions (*Upper*) and spectrographic displays (*Lower*) for the vowels /i/ as in "heat" and /a/ as in "hot." Formant frequencies, regions of the frequency spectrum in which the concentration of energy is high, are marked for each vowel.

A great deal has been learned since the debate ensued, caused largely by experiments conducted on infants. Infants' perception of the phonetic units of speech, which requires tracking the formant frequencies (Fig. 1) (5), and their detection of words from cues in running speech (Fig. 2) (6) support a different view. The emerging view argues that the kind of learning taking place in early language acquisition cannot be accounted for by Skinnerian reinforcement. On the other hand, the idea that language acquisition involves a selectionist process wherein language input operates on innately specified options also is not supported. The emerging view suggests that infants engage in a new kind of learning in which language input is mapped in detail by the infant brain. Six principles reflecting this view are offered.

*E-mail: pkkuhl@u.washington.edu.

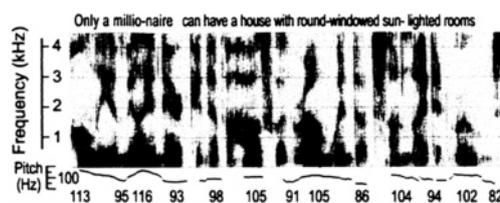


Fig. 2. Spectrographic display of running speech showing the formant frequencies and the pitch (fundamental frequency) of the voice over time. Increases in pitch indicate primary stress in the utterance. [Reproduced with permission from ref. 6 (Copyright 1999, Allyn & Bacon).]

Initial Perception Parses Speech Correctly and Is Universal, but Not Domain Specific or Species Specific

Any theory of language acquisition has to specify how infants parse the auditory world to make the critical units of language available. This is a formidable problem as indicated by the difficulty computers have in segmenting speech (7–9). Early experiments on infants confirmed their abilities to parse speech correctly at the phonetic level and revealed that their abilities are universal across languages. Interestingly, however, the data also demonstrated that the kind of partitioning seen for speech is not limited to humans or limited to speech.

The evidence derived from tests of categorical perception (10). When adult listeners were tested on a continuum that ranges from one syllable (such as “bat”) to another (“pat”), perception appeared absolute. Adults discriminated phonetic units that crossed the “phonetic boundary” between categories but not stimuli that fell within a category. The phenomenon was language-specific; Japanese adults, for example, failed to show a peak in discrimination at the phonetic boundary of an American English/ra-la/series (as in “rake” vs. “lake”) (11).

Categorical perception provided an opportunity to test whether infants could parse the basic units of language, and discrimination tests confirmed that they did. Infants discriminated only between stimuli from different phonetic categories (12–14). Moreover, unlike adults, infants demonstrated the effect for the phonetic units of all languages (15, 16). Eimas hypothesized that infants’ abilities reflected innate “phonetic feature detectors” that evolved for speech and theorized that infants are biologically endowed with neural mechanisms that respond to the phonetic contrasts used by the world’s languages (17).

Experimental tests on nonhuman animals altered this conclusion (18, 19). Animals also exhibited categorical perception; they demonstrated perceptual “boundaries” at locations where humans perceive a shift from one phonetic category to another (18, 19) (Fig. 3). In tests of discrimination, monkeys showed peaks in sensitivity that coincided with the phonetic boundaries used by languages (20–22) (Fig. 4). The results were subsequently replicated in a number of species (23, 24). Recently, additional tests on infants and monkeys revealed similarities in their perception of the prosodic cues of speech as well (25).

Two conclusions were drawn from the initial comparative work (26). First, infants’ parsing of the phonetic units at birth was a discriminative capacity that could be accounted for by a general auditory processing mechanism, rather than one that evolved specifically for speech. Differentiating the units of speech did not imply *a priori* knowledge of the phonetic units themselves, merely the capacity to detect differences between them, which was constrained in an interesting way (18, 19, 25, 27). Second, in the evolution of language, acoustic differences detected by the auditory perceptual processing mechanism strongly influenced the selection of phonetic units used in language. On this view, particular auditory features were exploited in the evolution of the sound system used in language (19, 26, 27). This ran counter to two prevailing principles at the time: (i) the view that phonetic units were prespecified in infants, and (ii) the view that language evolved in humans without continuity with lower species.

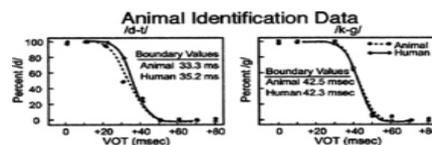


Fig. 3. Humans’ and animals’ identification functions for two series ranging from a voiced to a voiceless phonetic unit (/d/to/t/ and /g/to/k/) showing that animals’ perceptual boundaries correspond to humans’ phonetic boundaries. VOT, voice onset time. [Reproduced with permission from ref. 19 (Copyright 1978, Acoustical Society of America).]

Categorical perception also was demonstrated with nonspeech stimuli that mimicked speech features without being perceived as speech, in both adults (28, 29) and infants (30). This finding supported the view that domain-general mechanisms were responsible for infants’ initial partitioning of the phonetic units of language.

Development Is Not Based on Selection

Eimas’ early model of speech perception was selectionist in nature. An innate neural specification of all possible phonetic units allowed selection of a subset of those units to be triggered by language input (17). The notion was that linguistic experience produced either maintenance or loss. Detectors stimulated by ambient language were maintained, whereas those not stimulated by language input atrophied.

Developmental studies were initially seen as providing support for the selectionist view. Werker and her colleagues demonstrated that, by 12 months of age, infants no longer discriminate non-native phonetic contrasts, even though they did so at 6 months of age (31). The finding was interpreted as support for a selectionist theory; there was a “loss” of a subset of phonetic units initially specified.

Modifications regarding the extent to which listeners “lost” the ability to discriminate non-native phonetic units were quick to follow (32). Adult performance on non-native contrasts could be increased by a number of factors: (i) the use of techniques that minimize the effects of memory (33, 34), (ii) extensive training

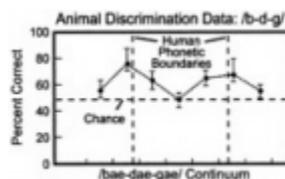


Fig. 4. Monkey discrimination performance for pairs of stimuli drawn from a continuum of speech sounds ranging from /b/to/d/to/g/, showing that sensitivity increases near the locations of humans’ phonetic boundaries. [Reproduced with permission from ref. 22 (Copyright 1978, Acoustical Society of America).]

(35, 36), and (iii) the use of contrasts, such as Zulu clicks, that are not related to native-language categories (37, 38). These data, indicating that there is not an immutable loss of phonetic abilities for non-native units (32), did not refute the selectionist position. The fact that listeners do not completely lose the ability to discriminate non-native contrasts does not alter the basic tenet of the selectionist view, which is that the role of language experience is to maintain or decrease the activity of innately specified neural detectors. To refute the selectionist position, studies must demonstrate that infants listening to ambient language are engaged in some other kind of learning process, a process that is not fundamentally subtractive in nature. New studies on learning provide that demonstration.

Infants' Learning Strategies "Map" Language Input

Learning theory as a mechanism for language acquisition had been dismissed by early theorists because of the failure of existing learning models, such as Skinner's, to explain the facts of language development (2). At present, however, learning models figure prominently in debates on language (39–42). What has changed? The discoveries of the last two decades, demonstrating that by simply listening to language infants acquire sophisticated information about its properties, have created new views of learning.

Three important examples of a new kind of learning have emerged. First, infants detect patterns in language input. Second, infants exploit the statistical properties of the input, enabling them to detect and use distributional and probabilistic information contained in ambient language to identify higher-order units. Third, infant perception is altered—literally warped—by experience to enhance language perception. No speaker of any language perceives acoustic reality; in each case, perception is altered in the service of language.

Infants Abstract Patterns. A major requirement of language processing is the detection of similarities, or patterns, in language input, a stumbling block for computer speech recognition (7). Infants demonstrate excellent skills at pattern recognition for speech. A number of studies have shown that 6-month-old infants, trained to produce a head-turn response when a sound from one category is presented (such as the vowel /a/ in "pop"), and to inhibit that response when an instance from another vowel category is presented (/i/ in "peep"), demonstrate the ability to perceptually sort novel instances into categories (43).

For example, infants perceptually sort vowels that vary across talkers and intonation contours (44, 45), as well as syllables that vary in their initial consonant (those beginning with /m/ as opposed to /n/, or those beginning with /s/ versus /ʃ/) across variations in talkers and vowel contexts (46, 47). Moreover, infants perceptually sort syllables based on a phonetic feature shared by their initial consonants, such as a set of nasal consonants, /m/, /n/, and /ŋ/, as opposed to a set of stop consonants, /b/, /d/, and /g/ (46). Recent tests show that 9-month-old infants are particularly attentive to the initial portions of syllables (48).

Infants' detection of patterns is not limited to phonetic units. More global prosodic patterns contained in language also are detected. At birth, infants have been shown to prefer the language spoken by their mothers during pregnancy, as opposed to another language (49–51). This skill requires infant learning of the stress and intonation pattern characteristic of the language (the pitch information shown in Fig. 2), information that is reliably transmitted through bone conduction to the womb (52). Additional evidence that the learning of speech patterns commences *in utero* stems from studies showing infant preference for their mother's voice over another female at birth (53) and their preference for stories read by the mother during the last 10 weeks of pregnancy (54).

Between 6 and 9 months, infants exploit prosodic patterns related to the stress or emphasis typical of words in their native language. In English, a strong/weak pattern of stress, with emphasis on the initial syllable ("baby" "mommy," "table") is typical, whereas a weak/strong pattern predominates in other languages. American infants tested at 6 months show no listening preference for words with the strong/weak as opposed to the weak/strong pattern, but by 9 months they exhibit a strong preference for the pattern typical of their native language (55). Infants also use prosodic cues to detect major constituent boundaries, such as clauses. At 4 months of age, infants listen equally long to Polish and English speech samples that have pauses inserted at clause boundaries as opposed to within clauses, but by 6 months, infants listen preferentially to pauses inserted at the clause boundaries appropriate only to their native language (41, 56).

By 9 months of age, infants detect patterns related to the orderings of phonemes that are legal for their language. In English, for example, the combination *zw* or *vl* is not legal; in Dutch, they are permissible. By 9 months of age, but not at 6 months of age, American infants listen longer to English words, whereas Dutch infants show a listening preference for Dutch words (57). At this age, infants do not recognize the words themselves, but recognize the perceptual patterns typical of words in their language. They develop a "perceptual sleeve" in which words fit; a description of word candidates assists them in identifying potential words in running speech.

Infants Exploit Statistical Properties of Language Input. Running speech presents a problem for infants because, unlike written speech, there are no breaks between words. New research shows that infants detect and exploit the statistical properties of the language they hear to find word candidates in running speech before they know the meanings of words. Goodsitt, Morgan, and Kuhl (58) demonstrated this in 7-month-old infants by using artificial words.

Goodsitt *et al.* examined infants' abilities to maintain the discrimination of two isolated syllables, /de/ and /ti/, when these target syllables were later embedded in three-syllable strings. The three-syllable strings contained the target syllable and a bisyllable composed of the syllables /ko/ and /ga/. The arrangement of /ko/ and /ga/ was manipulated to change the degree to which they could be perceived as a likely word candidate. Three conditions were tested. In *a*, /koga/ was an invariantly ordered "word," appearing either after the target syllables, /dekoga/ and /tikoga/, or before it, /kogade/ and /kogati/. In this condition, the transitional probability between the /ko/ and /ga/ was always 1.0. If infants detect /koga/ as a unit, it should assist infants in detecting and discriminating /de/ from /ti/. In *b*, the two syllables could either appear in variable order, either /koga/ or /gako/, reducing the transitional probabilities to 0.3 and preventing infants from perceiving /koga/ as a word. In *c*, one of the context syllables was repeated (e.g., /koko/). In this case, /koko/ could be perceived as a unit, but the basis of the perception would not be high transitional probabilities; the transitional probabilities between syllables in *c* remain low (0.3).

The results confirmed the hypothesis that 7-month-old infants exploit transitional probabilities. Infants discriminated the target syllables in condition *a* significantly more accurately than in either *b* or *c*, the latter of which showed equally poor discrimination. These strategies also have been shown to be effective for adults presented with artificial nonspeech analogs created by computer (42, 59).

In further work, Saffran, Aslin, and Newport (42) directly assessed 8-month-old infants' abilities to learn pseudowords based on transitional probabilities. Infants were exposed to 2-min strings of synthetic speech composed of four different

pseudowords that followed one another equally often. There were no breaks, pauses, stress differences, or intonation contours to aid infants in recovering these “words” from the strings of syllables. During the test phase, infants listened to two of the original pseudowords and two new words formed by combining parts of two of the original words. The results demonstrated that infants’ listened longer to the new words, demonstrating that they are capable of using statistical regularities to detect words (60).

Additional examples of the computation and use of probability statistics have been uncovered. Nine-month-old infants detect the probability of occurrence of legal sequences that occur in English (61). Certain combinations of two consonants are more likely to occur within words whereas others occur at the juncture between words. The combination “ft” is more common within words whereas the combination “vt” is more common between words. Nine-month-olds were tested with consonant (C) and vowel (V) strings of the form CVCCVC. These items contained embedded CCs that were either frequent or infrequent in English. Infants listened significantly longer to the lists containing frequent within-word CCs.

The results reveal that an old principle of Gestalt psychology, referred to as “common fate” (58), plays a role in speech perception. Phonemes that are typically linked, and thus share a common fate, are perceived as units by infants. It is interesting to note that early object perception also may rely on this principle. Physical entities whose properties cohere in space, and move together, are perceived as individuated objects (62). Whether the constraints underlying infants’ detection of common fate information for physical objects and speech are identical or different is important to theory and remains to be examined.

Language Experience Warps Perception. Language experience not only produces a change in infants’ discriminative abilities and listening preferences, it results in a “mapping” that alters perception. A research finding that helps explain this is called the perceptual magnet effect. The magnet effect is observed when tokens perceived as exceptionally good representatives of a phonetic category (“prototypes”) are used in tests of speech perception (63–66). Many behavioral (63–69) and brain (70–73) studies indicate that native-language phonetic prototypes evoke special responses when compared with nonprototypes.

When tested with a phonetic prototype as opposed to a nonprototype from the same category, infants show greater ability to generalize to other category members (63, 64). The prototype appears to function as a “magnet” for other stimuli in the category, in a way similar to that shown for prototypes of other cognitive categories (74, 75). Moreover, the perceptual magnet effect depends on exposure to a specific language (65). Six-month-old infants being raised in the United States and Sweden were tested with two vowel prototypes, an American English /i/ vowel prototype and a Swedish /y/ vowel prototype, using the exact same stimuli (Fig. 5A), techniques, and testers in the two countries. American infants demonstrated the magnet effect only for the American English /i/, treating the Swedish /y/ like a nonprototype. Swedish infants showed the opposite pattern, demonstrating the magnet effect for the Swedish /y/ and treating the American English /i/ as a nonprototype (Fig. 5B). The results show that by 6 months of age, perception is altered by language experience.

Categorical perception and the perceptual magnet effect make different predictions about the perception and organization underlying speech categories and appear to arise from different mechanisms (76). Interestingly, comparative tests show that, unlike categorical perception, animals do not exhibit the perceptual magnet effect (64).

In adults, the distortion of perception caused by language experience is well illustrated by a study on the perception of American English /r/ and /l/ in American and Japanese listeners. The /r-l/ distinction is difficult for Japanese speakers to perceive and produce; it is not used in the Japanese language (77, 78). In the study, Iverson and Kuhl (79) used computer-synthesized syllables beginning with /r/ and /l/, spacing them at equal physical intervals in a two-dimensional acoustic grid (Fig. 6A). American listeners identified each syllable as /r/ or /l/, rated its category goodness, and estimated the perceived similarity for all possible pairs of syllables. Similarity ratings were scaled by using multidimensional scaling techniques. The results provide a map of the perceived distances between stimuli—short distances for strong similarity and long distances for weak similarity. In the American map (Fig. 6B), magnet effects (seen as a shrinking of perceptual space) occur in the region of each category’s best instances. Boundary effects (seen as a stretching of perceptual space) occur at the division between the two categories.

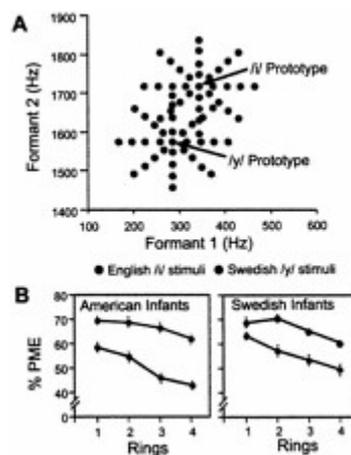


Fig. 5. (A) Formant frequencies of vowels surrounding an American/i/ prototype (red) and a Swedish/y/prototype (blue). (B) Results of tests on American and Swedish infants indicating an effect of linguistic experience. Infants showed greater generalization when tested with the native-language prototype. PME, Perceptual magnet effect. [Reproduced with permission from ref. 65 (Copyright 1992, American Association for the Advancement of Science).]

The experiment has recently been completed with Japanese monolingual listeners, and the results show a striking contrast in the way the /r-l/ stimuli are perceived by American and Japanese speakers. The map revealed by multidimensional scaling analysis is totally different—no magnet effects or boundary effects appear. Japanese listeners hear one category of sounds, not two, and attend to different dimensions of the same stimuli. The results suggest that linguistic experience produces mental maps for speech that differ substantially for speakers of different languages (40, 69, 79).

The important point regarding development is that the initial perceptual biases shown by infants in tests of categorical perception (12–16), as well as asymmetries in perception seen in infancy (80, 81), produce a contouring of the perceptual space that is universal. This universal contouring soon gives way to a language-specific mapping that distorts perception, completely revising the perceptual space underlying speech processing (65).

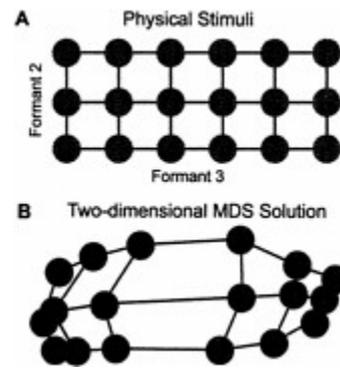


Fig. 6. (A) Physical distance between/ra-la/syllables in a grid created by varying formants 2 and 3 in equal steps. (B) Perceptual distance between syllables for American listeners showing a warping of acoustic space. MDS, Multidimensional scaling. [Reproduced with permission from ref. 79 (Copyright 1996, Acoustical Society of America).]

A model reflecting this developmental sequence from universal perception to language-specific perception, called the Native Language Magnet model, proposes that infants' mapping of ambient language warps the acoustic dimensions underlying speech, producing a complex network, or filter, through which language is perceived (39, 40, 82). The language-specific filter alters the dimensions of speech we attend to, stretching and shrinking acoustic space to highlight the differences between language categories. Once formed, language-specific filters make learning a second language much more difficult because the mapping appropriate for one's primary language is completely different from that required by other languages. Studies of adult bilinguals, who were exposed to their second language after the age of 6, demonstrate magnet effects only for the first language, illustrating the potent effects of early linguistic experience (66). According to the Native Language Magnet theory, infants' transition in speech perception between 6 and 12 months reflects the formation of a language-specific filter.

In summary, the studies on speech learning, demonstrating that infants detect patterns, extract statistical information, and have perceptual systems that can be altered by experience, cannot be explained by recourse to Skinnerian reinforcement learning. This is a different kind of learning, one ubiquitous during early development. Its study will be valuable beyond what it tells us about language learning.

Are the new learning strategies observed for speech domain-specific and/or species-specific? Research on cognitive development confirms the fact that categorization (83), statistical learning (84), and prototype effects (85) are not unique to speech. Further tests need to be done to determine the constraints operating on these abilities in infants by using linguistic and nonlinguistic events. What about animal tests? Thus far, the data suggest differences between animals and humans on these kinds of learning. For instance, monkeys do not exhibit the perceptual magnet effect (64). Animals do show some degree of internal structure for speech categories after extensive training (24), but it is unlikely the perceptual magnet effect would be spontaneously produced in an animal after 6 months' experience listening to language, as seen in human infants. Similarly, animals are sensitive to transitional probabilities (86), but it is unlikely that an animal would spontaneously exhibit statistical learning after simply listening to language, as human infants have been shown to do. These issues can be resolved with empirical tests.

Vocal Learning Unifies Perception and Production

Infants not only learn the perceptual characteristics of their language, they become native speakers, which requires imitation of the patterns of speech they hear others produce. Vocal learning critically depends on hearing the vocalizations of others and hearing oneself produce sound. This is true both for humans, who do not learn spoken language (or even babble normally) if they are deaf (87), and also for song birds (88). Production plays a role in normal language development; infants tracheostomized at the time at which they normally would babble show abnormal patterns of development that persist (89). These cases illustrate the strong dependency between perception and production and suggest why speech motor patterns learned early in life become difficult to alter later. Speakers who learn a second language after puberty produce it with an "accent" typical of their primary language, even after long-term instruction (90).

Imitation forges this early link between perception and production. By 1 year of age infants' spontaneous utterances reflect their imitation of ambient language patterns (91, 92). Laboratory studies indicate that the fundamental capacity to imitate sound patterns is in place even earlier. In a recent study, Kuhl and Meltzoff (93) recorded infant utterances at 12, 16, and 20 weeks of age while the infants watched and listened to a video recording of a woman producing a vowel, either /a/, /i/, or /u/ for 5 min on each of 3 successive days. The results demonstrate developmental change between 12 and 20 weeks—by 20 weeks, there is clear separation between the three vowel categories for infants (Fig. 7). At this age, infants clearly imitate the model, and their vowels have appropriate formant frequency values in relation to one another, even though infants' vowels occur in a much higher frequency range (93).

Early theories of speech perception held that speech was perceived with reference to production (10). The developmental data suggest a different conclusion—early in life, perceptual representations of speech are stored in memory. Subsequently, these representations guide the development of motor speech. The two systems are thus tightly coupled early on, but the coupling is seen as a coregistration of auditory and motor information, a polymodal mapping, rather than one in which the representation is specified in motor terms. Perceptual experience that guides sensory-motor learning also is seen in infants' imitation of nonspeech oral movements (94, 95) and in sign language (96). The perception-action links observed for speech thus may rely on domain-general capabilities.

In related studies, infants also show an ability to link mouth movements they see to auditory signals they hear. Studies on 18- to 20-week-old infants show that they look longer at a face pronouncing a vowel that matches one they hear as opposed to a mismatched face (97). Infants' polymodal speech representations are thus likely to contain information regarding visual, as well as auditory instantiations of speech (ref. 98, see also refs. 99–101).

“Motherese” Is Instructive

Historically, language input was seen as a trigger for selecting among innately specified options. New data suggest that language addressed to infants plays a much more important role. The universal speaking style used by caretakers around the world when they address infants, often called “motherese” or “parentese” (102), has been shown to be preferred over adult-directed speech by infants given a choice (103, 104). Moreover, the exaggerated stress and increased pitch typical of infant-directed speech assists infants in discriminating phonetic units (105).

Infant-directed speech also is altered at the phonetic level and these alterations are argued to help infants learn. In a recent

study, women were recorded while speaking to their 2-month-old infants and to another adult in the United States, Russia, and Sweden (106). Mothers used the vowels /i/, /a/, and /u/, in both settings, and their speech was analyzed spectrographically. The results demonstrated that the phonetic units of infant-directed speech are acoustically exaggerated. The results show a stretching of the acoustic space encompassing speech (Fig. 8). Exaggerating speech not only makes it more discriminable for infants, it highlights critical parameters used in the native language. This may aid infants' discovery of the dimensions of sound used in their native language. Mothers addressing infants also increase the variety of exemplars they use, behaving in a way that makes mothers resemble many different talkers, a feature shown to assist category learning in second-language learners (107). In recent studies, language-delayed children show substantial improvements in measures of speech and language after listening to speech altered by computer to exaggerate phonetic differences (108, 109).

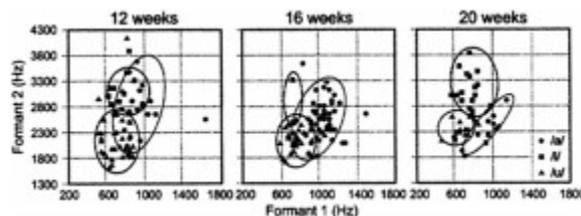


Fig. 7. Infants' vowels recorded as they imitate an adult show developmental change between 12 and 20 weeks of age. [Reproduced with permission from ref. 93 (Copyright 1996, Acoustical Society of America).]

Mothers addressing infants make other adjustments that appear to aid learning. When introducing new words, parents repeat the word often in stereotyped frames ("Where's the __," "See the __," "That's a __") (110), which would highlight the items in sentence-final position. They also present new words in a great variety of contexts, which would highlight the internal transitional probabilities of the new words against the backdrop of a variety of contexts (58). These new data suggest that the modifications made by adults unconsciously when they speak to infants plays a role in helping infants map native-language input. This represents a change in theoretical perspective with regard to the role of motherese in language acquisition.

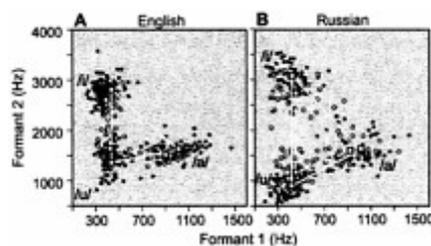


Fig. 8. Formant frequencies for vowels produced by American (A) and Russian (B) mothers as they talked to their infants (solid symbols) and to another adult (open symbols), indicating that vowels in infant-directed speech are acoustically exaggerated. [Reproduced with permission from ref. 106 (Copyright 1997, American Association for the Advancement of Science).]

The Critical Period for Language Learning Depends on Experience, Not Just Time

There is no doubt that children learn language more naturally and efficiently than adults, a paradox given adults' superior cognitive skills. The question is: Why?

Language acquisition often is cited as an example of a "critical period" in development, a learning process that is constrained by time, or factors such as hormones, that are outside the learning process itself. The studies on speech suggest an alternative (40, 82). The work on speech suggests that later learning may be constrained by the initial mapping that has taken place. For instance, if learning involves the creation of mental maps for speech, as suggested by the Native Language Magnet model (65, 82), it likely "commits" neural structure in some way. Measurements of brain activity, for example, confirm left-hemisphere effects for native-language sounds in the mismatched negativity (MMN), an event-related potential elicited by a change in a repetitive sound pattern (72). In infants, the MMN is observed to changes in both native and nonnative contrasts at 6 months of age. At 12 months of age, the MMN exists only for native language contrasts (73). Neural commitment to a learned structure may interfere with the processing of information that does not conform to the learned pattern. On this account, initial learning can alter future learning independent of a strictly timed period.

Support for the neural commitment view comes from two sources, second language learning, and training studies. When acquiring a second language, certain phonetic distinctions are notoriously difficult to master both in speech perception and production, as shown, for example, by the difficulty of the /r-/l/ distinction for native speakers of Japanese, even after training (11, 78, 111, 112). The hypothesis is that, for Japanese people, learning to process English requires the development of a new map, one more appropriate for English. New training studies suggest that exaggerating the dimensions of foreign language contrasts (36), as well as providing listeners with multiple instances spoken by many talkers (113), are effective training methods. These studies show that feedback and reinforcement are not necessary in this process; listeners simply need the right kind of listening experience (36, 113). Interestingly, the features shown to assist second-language learners—exaggerated acoustic cues, multiple instances by many talkers, and mass listening experience—are features that motherese provides infants.

Early in life, interference effects are minimal and two different mappings can be acquired, as is the case for infants learning two languages. Anecdotal evidence suggests that infants exposed to two languages do much better if each parent speaks one of the two languages, rather than both parents speaking both languages. This may be the case because it is easier to map two different sets of phonetic categories (one for each of the two languages) if they can be perceptually separated. A second

language learned later in life (after puberty) may require another form of separation between the two systems to avoid interference. Data gathered by using functional MRI techniques indicate that adult bilinguals who acquire both languages early in life activate overlapping regions of the brain when processing the two languages, whereas those who learn the second language later in life activate two distinct regions of the brain for the two languages (114). This is consistent with the idea that the brain's processing of a primary language can interfere with the second language. The problem is avoided if both are learned early in development.

Conclusions

The framework that emerges from this research is very different from that held historically. Infants are neither the tabula rasas that Skinner described nor the innate grammarians that Chomsky envisioned. Infants have inherent perceptual biases that segment phonetic units without providing innate descriptions of them. They use inherent learning strategies that were not expected, ones thought to be too complex and difficult for infants to use. Adults addressing infants unconsciously modify speech in ways that assist the brain mapping of language. In combination, these factors provide a powerful discovery procedure for language. Six tenets of a new view of language acquisition are offered: (i) infants' initially parse the basic units of speech allowing them to acquire higher-order units created by their combinations; (ii) the developmental process is not a selectionist one in which innately specified options are selected on the basis of experience; (iii) rather, a perceptual learning process, unrelated to Skinnerian learning, commences with exposure to language, during which infants detect patterns, exploit statistical properties, and are perceptually altered by that experience; (iv) vocal imitation links speech perception and production early, and auditory, visual, and motor information are coregistered for speech categories; (v) adults addressing infants unconsciously alter their speech to match infants' learning strategies, and this is instrumental in supporting infants' initial mapping of speech; and (vi) the critical period for language is influenced not only by time, but by the neural commitment that results from experience.

Taken together, these principles suggest that what is innate regarding language is not a universal grammar and phonetics, but innate biases and strategies that place constraints on perception and learning. They allow infants to recover from language input the rules by which people in their community communicate. Language is thus innately discoverable, but not innate in the way that selectionist models suggested. The learning strategies used by infants may themselves have influenced the nature of language, in much the same way that general auditory processing influenced the selection of phonetic units for language during its evolution. The continued study of language development by infants promises to reveal the precise nature of the relationship between language and mind.

I thank Erica Stevens and Feng-Ming Tsao for assistance on preparation of the manuscript and Andy Meltzoff for comments on the issues discussed. The preparation of this manuscript and my research are supported by grants from the National Institutes of Health (HD37954) and the Human Frontiers Science Program (RG0159).

1. Skinner, B.F. (1957) *Verbal Behavior* (Appleton-Century-Crofts, New York).
2. Chomsky, N. (1957) *Language* **35**, 26–58.
3. Wexler, K. & Culicover, P.W. (1980) *Formal Principles of Language Acquisition* (MIT Press, Cambridge, MA).
4. Fodor, J.A. (1983) *The Modularity of Mind: An Essay on Faculty Psychology* (MIT Press, Cambridge, MA).
5. Stevens, K.N. (1998) *Acoustic Phonetics* (MIT Press, Cambridge, MA).
6. Pickett, J.M. (1999) *The Acoustics of Speech Communication* (Allyn and Bacon, Boston).
7. Goss, N., Judge, P.C., Port, O. & Wildstrom, S.H. (1998) *BusinessWeek* February **23**, 60–72.
8. Waibel, A. (1986) in *Pattern Recognition by Humans and Machines*, eds. Schwab, E.C. & Nusbaum, H.C. (Academic, New York), pp. 159–186.
9. Bernstein, J. & Franco, H. (1996) in *Principles of Experimental Phonetics*, ed. Lass, N.J. (Mosby, St. Louis), pp. 408–434.
10. Liberman, A.M., Cooper, F.S., Shankweiler, D.P. & Studdert-Kennedy, M. (1967) *Psychol. Rev.* **74**, 431–461.
11. Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Jenkins, J.J. & Fujimura, O. (1975) *Percept. Psychophys.* **18**, 331–340.
12. Eimas, P.D., Siqueland, E.R., Jusczyk, P. & Vigorito, J. (1971) *Science* **171**, 303–306.
13. Eimas, P.D. (1974) *Percept. Psychophys.* **16**, 513–521.
14. Eimas, P.D. (1975) *Percept. Psychophys.* **18**, 341–347.
15. Lasky, R.E., Syrdal-Lasky, A. & Klein, R.E. (1975) *J. Exp. Child Psych.* **20**, 215–225.
16. Streeter, L.A. (1976) *Nature (London)* **259**, 39–41.
17. Eimas, P.D. (1975) in *Infant Perception: Vol. 2. From Sensation to Cognition*, eds. Cohen, L.B. & Salapatek, P. (Academic, New York), pp. 193–231.
18. Kuhl, P.K. & Miller, J.D. (1975) *Science* **190**, 69–72.
19. Kuhl, P.K. & Miller, J.D. (1978) *J. Acoust. Soc. Am.* **63**, 905–917.
20. Kuhl, P.K. (1981) *J. Acoust. Soc. Am.* **70**, 340–349.
21. Kuhl, P.K. & Padden, D.M. (1982) *Percept. Psychophys.* **32**, 542–550.
22. Kuhl, P.K. & Padden, D.M. (1983) *J. Acoust. Soc. Am.* **73**, 1003–1010.
23. Dooling, R.J., Best, C.T. & Brown, S.D. (1995) *J. Acoust. Soc. Am.* **97**, 1839–1846.
24. Kluender, K.R., Diehl, R.L. & Killeen, P.R. (1987) *Science* **237**, 1195–1197.
25. Ramus, F., Hauser, M.D., Miller, C., Morris, D. & Mehler, J. (2000) *Science* **288**, 349–351.
26. Kuhl, P.K. (1991) in *Plasticity of Development*, eds. Brauth, S.E., Hall, W.S. & Dooling, R.J. (MIT Press, Cambridge, MA), pp. 73–106.
27. Kuhl, P.K. (1988) *Hum. Evol.* **3**, 19–43.
28. Miller, J.D., Wier, C.C., Pastore, R.E., Kelly, W.J. & Dooling, R.J. (1976) *J. Acoust. Soc. Am.* **60**, 410–417.
29. Pisoni, D.B. (1977) *J. Acoust. Soc. Am.* **61**, 1352–1361.
30. Jusczyk, P.W., Rosner, B.S., Cutting, J.E., Foard, C.F. & Smith, L.B. (1977) *Percept. Psychophys.* **21**, 50–54.
31. Werker, J.F. & Tees, R.C. (1984) *Inf. Behav. Dev.* **7**, 49–63.
32. Werker, J.F. (1995) in *An Invitation to Cognitive Science: Language*, eds. Gleitman, L.R. & Liberman, M. (MIT Press, Cambridge, MA), pp. 87–107.
33. Werker, J.F. & Logan, J.S. (1985) *Percept. Psychophys.* **37**, 35–44.
34. Carney, A.E., Widin, G.P. & Viemeister, N.F. (1977) *J. Acoust. Soc. Am.* **62**, 961–970.
35. Logan, J.S., Lively, S.E. & Pisoni, D.B. (1991) *J. Acoust. Soc. Am.* **89**, 874–886.
36. McClelland, J.L., Thomas, A., McCandliss, B.D. & Fiez, J.A. (1999) in *Brain, Behavioral, and Cognitive Disorders: The Neurocomputational Perspective*, eds. Reggia, J., Ruppin, E. & Glanzman, D. (Elsevier, Oxford), pp. 75–80.
37. Best, C.T., McRoberts, G.W. & Sithole, N.M. (1988) *J. Exp. Psych. Hum. Percept. Perform.* **14**, 345–360.
38. Best, C.T. (1995) in *Advances in Infancy Research*, eds. Rovee-Collier, C. & Lipsitt, L.P. (Ablex, Norwood, NJ), pp. 217–304.
39. Kuhl, P.K. (1994) *Curr. Opin. Neurobiol.* **4**, 812–822.
40. Kuhl, P.K. (2000) in *The New Cognitive Neurosciences*, ed. Gazzaniga, M.S. (MIT Press, Cambridge, MA), 2nd Ed., pp. 99–115.
41. Jusczyk, P.W. (1997) *The Discovery of Spoken Language* (MIT Press, Cambridge, MA).
42. Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996) *Science* **274**, 1926–1928.
43. Kuhl, P.K. (1985) in *Neonate Cognition: Beyond the Blooming Buzzing Confusion*, eds. Mehler, J. & Fox, R. (Erlbaum, Hillsdale, NJ), pp. 231–262.
44. Kuhl, P.K. (1979) *J. Acoust. Soc. Am.* **66**, 1668–1679.
45. Kuhl, P.K. (1983) *Inf. Behav. Dev.* **6**, 263–285.
46. Hillenbrand, J. (1983) *J. Speech Hear. Res.* **26**, 268–282.
47. Kuhl, P.K. (1980) in *Child Phonology: Vol. 2. Perception*, eds. Yeni-Komshian, G.H., Kavanagh, J.F. & Ferguson, C.A. (Academic, New York), pp. 41–66.
48. Jusczyk, P.W. (1999) *Trends Cognit. Sci.* **3**, 323–328.
49. Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J. & Amiel-Tison, C. (1988) *Cognition* **29**, 143–178.
50. Moon, C., Cooper, R.P. & Fifer, W.P. (1993) *Inf. Behav. Dev.* **16**, 495–500.
51. Nazzi, T., Bertoni, J. & Mehler, J. (1998) *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 756–766.
52. Lecanuet, J.P. & Granier-Deferre, C. (1993) in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, eds. de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., McNeilage, P. & Morton, J. (Kluwer, Dordrecht, The Netherlands).
53. DeCasper, A.J. & Fifer, W.P. (1980) *Science* **208**, 1174–1176.

54. DeCasper, A.J. & Spence, M.J. (1986) *Inf. Behav. Dev.* **9**, 133–150.
55. Jusczyk, P.W., Cutler, A. & Redanz, N.J. (1993) *Child Dev.* **64**, 675–687.
56. Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Cassidy, K.W., Druss, B. & Kennedy, L. (1987) *Cognition* **26**, 269–286.
57. Jusczyk, P.W., Friederici, A.D., Wessels, J.M.L., Svenkerud, V.Y. & Jusczyk, A.M. (1993) *J. Mem. Lang.* **32**, 402–420.
58. Goodsitt, J.V., Morgan, J.L. & Kuhl, P.K. (1993) *J. Child Lang.* **20**, 229–252.
59. Wolff, J.G. (1977) *Br. J. Psych.* **68**, 97–106.
60. Aslin, R.N., Saffran, J.R. & Newport, E.L. (1998) *Psychol. Sci.* **9**, 321–324.
61. Mattys, S.L., Jusczyk, P.W., Luce, P.A. & Morgan, J.L. (1999) *Cog. Psych.* **38**, 465–494.
62. Spelke, E. (1994) *Cognition* **50**, 431–445.
63. Grieser, D. & Kuhl, P.K. (1989) *Dev. Psych.* **25**, 577–588.
64. Kuhl, P.K. (1991) *Percept. Psychophys.* **50**, 93–107.
65. Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N. & Lindblom, B. (1992) *Science* **255**, 606–608.
66. Bosch, L., Costa, A. & Sebastian-Galles, N. (2000) *Eur. J. Cognit. Psychol.* **12**, 189–221.
67. Samuel, A.G. (1982) *Percept. Psychophys.* **31**, 307–314.
68. Miller, J.L. (1994) *Cognition* **50**, 271–285.
69. Iverson, P. & Kuhl, P.K. (1995) *J. Acoust. Soc. Am.* **97**, 553–562.
70. Aaltonen, O., Eerola, O., Hellström, A., Uusipaikka, E. & Lang, A.H. (1997) *J. Acoust. Soc. Am.* **101**, 1090–1105.
71. Sharma, A. & Dorman, M.F. (1998) *J. Acoust. Soc. Am.* **104**, 511–517.
72. Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., *et al.* (1997) *Nature (London)* **385**, 432–434.
73. Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O. & Näätänen, R. (1995) *Hear. Res.* **82**, 53–58.
74. Medin, D.L. & Barsalou, L.W. (1987) in *Categorical Perception: The Groundwork of Cognition*, ed. Harnad, S. (Cambridge Univ. Press, New York), pp. 455–490.
75. Mervis, C.B. & Rosch, E. (1981) *Annu. Rev. Psychol.* **32**, 89–115.
76. Iverson, P. & Kuhl, P.K. (2000) *Percept. Psychophys.* **62**, 874–886.
77. Strange, W. & Dittmann, S. (1984) *Percept. Psychophys.* **36**, 131–145.
78. Goto, H. (1971) *Neuropsychologia* **9**, 317–323.
79. Iverson, P. & Kuhl, P.K. (1996) *J. Acoust. Soc. Am.* **99**, 1130–1140.
80. Polka, L. & Bohn, O.S. (1996) *J. Acoust. Soc. Am.* **100**, 577–592.
81. Miller, J.L. & Eimas, P.D. (1996) *Percept. Psychophys.* **58**, 1157–1167.
82. Kuhl, P.K. (1998) in *Mechanistic Relationships Between Development and Learning*, eds. Carew, T.J., Menzel, R. & Shatz, C.J. (Wiley, New York), pp. 53–73.
83. Younger, B.A. & Cohen, L.B. (1985) in *The Psychology of Learning and Motivation*, ed. Bower, G.H. (Academic, San Diego), Vol. 19, pp. 211–247.
84. Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999) *Cognition* **70**, 27–52.
85. Quinn, P.C. & Eimas, P.D. (1998) *J. Exp. Child Psychol.* **69**, 151–174.
86. Gallistel, C.R. (1990) *The Organization of Learning* (MIT Press, Cambridge, MA).
87. Oller, D.K. & MacNeilage, P.F. (1983) in *The Production of Speech*, ed. MacNeilage, P.F. (Springer, New York), pp. 91–108.
88. Doupe, A. & Kuhl, P.K. (1999) *Annu. Rev. Neurosci.* **22**, 567–631.
89. Locke, J.L. & Pearson, D.M. (1990) *J. Child Lang.* **17**, 1–16.
90. Flege, J.E. (1993) *J. Acoust. Soc. Am.* **93**, 1589–1608.
91. de Boysson-Bardies, B. (1993) in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, eds. de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., McNeilage, P. & Morton, J. (Kluwer, Dordrecht, The Netherlands), pp. 353–363.
92. Vihman, M.M. & de Boysson-Bardies, B. (1994) *Phonetica* **51**, 159–169.
93. Kuhl, P.K. & Meltzoff, A.N. (1996) *J. Acoust. Soc. Am.* **100**, 2425–2438.
94. Meltzoff, A.N. & Moore, M.K. (1977) *Early Dev. Parent.* **6**, 179–192.
95. Meltzoff, A.N. & Moore, M.K. (1994) *Inf. Behav. Dev.* **17**, 83–99.
96. Petitto, L.A. & Marentette, P.F. (1991) *Science* **251**, 1493–1496.
97. Kuhl, P.K. & Meltzoff, A.N. (1982) *Science* **218**, 1138–1141.
98. Kuhl, P.K. & Meltzoff, A.N. (1997) in *The Inheritance and Innateness of Grammars*, ed. Gopnik, M. (Oxford Univ. Press, New York), pp. 7–44.
99. MacKain, K., Studdert-Kennedy, M., Spieker, S. & Stern, D. (1983) *Science* **219**, 1347–1349.
100. Rosenblum, L.D., Schmuckler, M.A. & Johnson, J.A. (1997) *Percept. Psychophys.* **59**, 347–357.
101. Walton, G.E. & Bower, T.G.R. (1993) *Inf. Behav. Dev.* **16**, 233–243.
102. Ferguson, C.A. (1964) *Am. Anthropol.* **66**, 103–114.
103. Fernald, A. (1985) *Inf. Behav. Dev.* **8**, 181–195.
104. Fernald, A. & Kuhl, P. (1987) *Inf. Behav. Dev.* **10**, 279–293.
105. Karzon, R.G. (1985) *J. Exp. Child Psych.* **39**, 326–342.
106. Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevni-kova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U. & Lacerda, F. (1997) *Science* **277**, 684–686.
107. Lively, S.E., Logan, J.S. & Pisoni, D.B. (1993) *J. Acoust. Soc. Am.* **94**, 1242–1255.
108. Merzenich, M.M., Jenkins, W.M., Johnston, P., Schreiner, C., Miller, S.L. & Tallal, P. (1996) *Science* **271**, 77–81.
109. Tallal, P., Miller, S.L., Bedi, G., Byrna, G., Wang, X., Nagarajan, S.S., Schreiner, C., Jenkins, W.M. & Merzenich, M.M. (1996) *Science* **271**, 81–84.
110. Peters, A.M. (1983) *The Units of Language Acquisition* (Cambridge Univ. Press, Cambridge).
111. Flege, J.E., Takagi, N. & Mann, V. (1995) *Lang. Speech* **38**, 25–55.
112. Yamada, R.A. & Tohkura, Y. (1992) *Percept. Psychophys.* **52**, 376–392.
113. Pisoni, D.B. (1992) in *Speech Perception, Production and Linguistic Structure*, eds. Tohkura, Y., Vatikiotis-Bateson, E. & Sagisaka, Y. (Ohmsha, Tokyo), pp. 143–151.
114. Kim, K.H.S., Relkin, N.R., Lee, K.M. & Hirsch, J. (1997) *Nature (London)* **388**, 172–174.

NATIONAL ACADEMY OF SCIENCES COLLOQUIUM On Auditory Neuroscience: Development, Transduction, and Integration

19–21 May 2000

Arnold and Mabel Beckman Center and the Hyatt Newporter Hotel
Irvine, California

Poster Presentations

1. “Expression of GATA3 and Pax2 during development of the mouse inner ear”
Authors: Lawoko G., M.N.Rivolta and M.C.Holley
Department of Physiology, University of Bristol, United Kingdom
2. “Gene duplications gave rise to an ear-specific claudin in the zebrafish”
Authors: Kollmar, R., S.Karia, J.A.Kappler, and A.J.Hudspeth
Laboratory of Sensory Neuroscience, The Rockefeller University, New York, New York
3. “Modelling the effects of lateral inhibition in the developing chick basilar papilla”
Author: Pickles, J.O.
Vision Touch and Hearing Research Centre, University of Queensland, Queensland, Australia
4. “Primordial rhythmic bursting in embryonic cochlear ganglion cells”
Authors: Jones, S.M. and T.A.Jones
Department of Surgery, Physiology and ENT, University of Missouri School of Medicine, Columbia, Missouri
5. “An empirical model for signal processing by type II vestibular hair-cell/afferent synaptic adaptation”
Authors: Rabbitt, R.D., R.Boyle and S.M.Highstein
Department of Bioengineering, University of Utah, Salt Lake City, Utah, Center for Bioinformatics, Ames Research Center, NASA, Moffett Field, California, and Department of Otolaryngology, Washington University School of Medicine, St. Louis, Missouri
6. “Single unit responses in the dorsal cochlear nucleus of the Purkinje cell degeneration mice and control mice”
Authors: Parham, K. and D.O.Kim
Division of Otolaryngology, Department of Surgery and Department of Neuroscience, University of Connecticut Health Center, Farmington, Connecticut
7. “Psychophysical estimation of cochlear phase response: Level effects”
Authors: Leek, M.R., J.J.Lentz, and L.E.Dreisbach
Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, D.C.
8. “Effect of auditory cortex lesions on the discrimination of frequency change”
Authors: Harrington, I.A. and H.E.Heffner
Department of Psychology, University of Toledo, Toledo, Ohio

9. "Auditory neuropathy: "I can hear but do not understand"
Authors: Zeng, F.-G., A.Starr and S.Oba
Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland, University of California, Irvine, CA, and House Ear Institute, Los Angeles, California
10. "Frequency tuning plasticity in rat primary auditory cortex"
Authors: Pandya, P.K., J.Vazquez, D.Rathbun, N.D.Engineer, R.Moucha, and M.P. Kilgard
Cortical Plasticity Laboratory, Cognition and Neuroscience Program, University of Texas at Dallas, Richardson, Texas
11. "Temporal integration in the auditory cortex of awake primates"
Authors: Lu, T., L.Liang and X.Wang
Department of Biomedical Engineering, The Johns Hopkins University, Baltimore, Maryland
12. "Sensitivity of auditory cortical neurons to coherence of amplitude and frequency modulation"
Authors: Barbour, D.L. and X.Wang
Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland
13. "Song-selectivity in the zebra finch song nucleus HVc: Contributions of local and extrinsic inputs"
Authors: Rosen, M. and R.Mooney
Department of Neurobiology, Duke University, Durham, North Carolina
14. "Harmonic structure in the mammalian auditory cortex"
Authors: Kadia, S.C., D.K.Ryugol and X.Wang
Center for Hearing Sciences, Johns Hopkins University, Baltimore, Maryland

NATIONAL ACADEMY OF SCIENCES COLLOQUIUM On Auditory Neuroscience: Development, Transduction, and Integration

19–21 May 2000

Arnold and Mabel Beckman Center and the Hyatt Newporter Hotel
Irvine, California

May 19, 2000

Development and Regeneration of the Inner Ear

Julian Lewis, University College London, England

“The development of sensory patches: Hearing lessons from the fly”

Donna M.Fekete, Purdue University

“Molecular genetics of pattern formation in the developing inner ear”

Doris K.Wu, National Institutes of Health

“Shaping of the mammalian cochlea”

Richard A.Baird, Washington University School of Medicine

“Hair-cell repair and regeneration in normal and mitotically-blocked saccular cultures”

Jeffrey T.Corwin, University of Virginia School of Medicine

“Arousing the progenitors of mammalian hair cells”

David P.Corey, Harvard Medical School

“Sensory adaptation by hair cells”

May 20, 2000

Transduction, Adaptation, and Mechanical Amplification by Hair Cells

Geoffrey A.Manley, Technische Universität München, Germany

“Cochlear mechanisms from a phylogenetic viewpoint”

Mario A.Ruggero, Northwestern University

“Mechanics of the mammalian cochlea”

Ian J.Russell, University of Sussex, England

“Spatial and temporal representation of a tone on the guinea pig basilar membrane”

Jonathan F.Ashmore, University College London, England

“Molecular mechanisms in sound amplification in the mammalian cochlea”

A.J.Hudspeth, Rockefeller University

“Mechanical amplification by hair bundles”

May 20, 2000

Poster Session (Lecture Room)

Coding in the Auditory Central Nervous System

Donata Oertel, University of Wisconsin—Madison

“Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus”

Eric D.Young, Johns Hopkins University School of Medicine

“What does a receptive field mean? Insights from the dorsal cochlear nucleus”

Masakazu Konishi, California Institute of Technology

“Cellular mechanisms for resolving phase ambiguity in the owl’s inferior colliculus”

Jon H.Kaas, Vanderbilt University

“Subdivisions of auditory cortex and processing streams in primates”

Josef P.Rauschecker, Georgetown University Medical Center

“Parallel processing in the auditory cortex of human and nonhuman primates”

Nobuo Suga, Washington University

“The corticofugal system for hearing”

May 21, 2000

Coding and Plasticity in the Auditory Central Nervous System

Eric I.Knudsen, Stanford University School of Medicine

“Mechanisms of adaptive plasticity in the owl’s auditory localization pathway”

Andrew J.King, Oxford University, England

“Plasticity in the neural coding of auditory space in the mammalian brain”

Gregg H.Recanzone, University of California, Davis

“Cortical processing of sound location and temporal order”

Coffee break

Allison J.Doupe, University of California, San Francisco

“Neural learning of birdsong”

Xiaoqin Wang, Johns Hopkins University School of Medicine

“Cortical mechanisms underlying vocal communication in primates”

Patricia Kuhl, University of Washington

“Early experience and language development: Reinterpreting the critical period”

Adjournment