Optical Approaches for Neuroprosthetic Stimulation and Chronic Visualization of Sound Processing in the Adult Auditory System

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Hight, Ariel. 2019. Optical Approaches for Neuroprosthetic Stimulation and Chronic Visualization of Sound Processing in the Adult Auditory System. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:42013137">http://nrs.harvard.edu/urn-3:HUL.InstRepos:42013137</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Optical approaches for neuroprosthetic stimulation and chronic visualization of sound processing in the adult auditory system

A dissertation presented

by

Ariel Edward Hight

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Speech and Hearing Bioscience and Technology

Harvard University
Cambridge, Massachusetts
July 2019
Optical approaches for neuroprosthetic stimulation and chronic visualization of sound processing in the adult auditory system

ABSTRACT

Disabling hearing loss affects approximately 433 million people world-wide, resulting in impaired speech comprehension and associated hearing disorders including tinnitus and hyperacusis. Auditory multi-channel neuroprosthetic devices are widely used bioengineering solutions for sensory restoration. Auditory neuroprosthetics bypass sensory cells or primary neurons in the inner ear to directly stimulate early stages of the auditory pathway. These devices reliably restore the perception of auditory environmental awareness but have a mixed record of success in restoring unaided speech cues that form the basis of inter-personal communication. This dissertation uses rodent models to explore new approaches for stimulating the central auditory pathway and new approaches to understand how the auditory regions of the brain reorganize after hearing loss. In chapters 1-3, we describe efforts to identify neural bottlenecks in providing usable auditory cues through electrical stimulation of the brainstem and describe efforts to improve the spatial resolution of stimulation through optogenetic stimulation. In chapter 4, we describe a new approach for studying the organization of sound frequency representations in the auditory cortex using chronic widefield epifluorescence imaging of genetically encoded calcium indicators. In chapter 5, we leverage this long-term imaging approach to study the dynamics of compensatory plasticity following sudden acoustic trauma and associate these cortical dynamics with commensurate behavioral changes in loudness perception. Taken together, these studies suggest that optical stimulation of genetically encoded opsins
represents a promising candidate for improving the spatial specificity of auditory
neuroprostheses. On the flip side of the coin, optical approaches to track daily dynamics in sound
processing across large brain regions with high resolution will provide deeper insight into the
downstream processing of sound features after sensorineural hearing loss. The dissertation
discusses how neuroprosthetic stimulation strategies guided by the abstracted modes of sound
processing observed later in the auditory pathway rather than low-level auditory biophysics
could introduce a new chapter in the implementation of sensory restoration devices.
Table of Contents

ABSTRACT .................................................................................................................................................. III

ACKNOWLEDGMENTS .......................................................................................................................... VI

ACKNOWLEDGEMENT OF PUBLISHED WORK .................................................................................. IX

LIST OF TABLES AND FIGURES ......................................................................................................... X

CHAPTER 1 INTRODUCTION ................................................................................................................ 1

CHAPTER 2 AUDITORY BRAINSTEM STIMULATION WITH A CONFORMABLE MICROFABRICATED ARRAY ELICITS RESPONSES WITH TONO TOPICALLY ORGANIZED COMPONENTS ......................... 5

CHAPTER 3 SUPERIOR TEMPORAL RESOLUTION OF CHRONOS VERSUS CHANNELRHODOPSIN-2 IN AN OPTOGENETIC MODEL OF THE AUDITORY BRAINSTEM IMPLANT ............................................. 40

CHAPTER 4 HEARING THE LIGHT: NEURAL AND PERCEPTUAL ENCODING OF OPTOGENETIC STIMULATION IN THE CENTRAL AUDITORY PATHWAY ........................................................................ 58

CHAPTER 5 CELLULAR AND WIDEFIELD IMAGING OF SOUND FREQUENCY ORGANIZATION IN PRIMARY AND HIGHER-ORDER FIELDS OF THE MOUSE AUDITORY CORTEX ..................................... 86

CHAPTER 6 CORTICAL MAP DYNAMICS ASSOCIATED WITH PERCEPTUAL HYPERSENSITIVITY FOLLOWING ACOUSTIC TRAUMA .................................................................................................................. 134

CHAPTER 7 CONCLUSIONS .................................................................................................................. 164

REFERENCES .............................................................................................................................................. 165
ACKNOWLEDGMENTS

I have much to be thankful in my professional and personal life that have helped make this the completion of this report possible.

First and foremost, I would like to acknowledge my mentors, Daniel B Polley (Dan) and M. Christian Brown (Chris). Having the opportunities to work in their laboratories have been by itself a blessing, everything else has been icing on the cake. While I started my studies with Chris, his support during my time in the auditory brainstem implant lab was key for my growing confidence to bridge disciplines between the engineering and neuroscience realms. Moreover, his continual mentorship in the lab and facilitation of many social activities of the lab throughout the duration of my studies are immensely appreciated and have helped make me a better-rounded and happier person. I have a strong desire to convey my appreciation to Dan for taking me under his wing as a 3rd/4th year student and supporting me through to the end of my studies. The active mentorship and deliberately thought out projects turned out to be a good fit for both my skill-set and in addressing my gaps in knowledge and expertise. Thank you also for leading a lab that has been incredibly engaging and never boring! I’d also like to acknowledge Dan Lee, a co-PI in the ABI lab, for providing and supporting me with the freedom to explore projects in the lab, which turned out to be necessary for identifying some of the interesting work produced in this report, and for selflessly sharing his wisdom as a clinician.

The interdisciplinary nature of the projects included in this dissertation required many collaborations. While the nature of collaborative work provides opportunities for investigating previously uncharted territory, collaborative work also requires carefully coordinated teamwork and I consider myself very fortunate to work with some wonderful people. First, I’d like to acknowledge my co-authors and a few people in particular. Amélie Guex and I had many
exciting and engaging conversations in my early years about ideas of how to study and improve the ABI. Another collaborator in my early years, Elliott Kozin, whom I share co-first authorship on the Chronos vs ChR2 paper, was generous in sharing his wisdom as a more senior researcher and as a clinician and was a wonderful triathlon buddy. Another collaborator in my early years was Wei Guo, who was also very generous with his time and in sharing his sharp humor which made working a bit more enjoyable. Towards the end of my studies in the ABI lab, I started on an interesting but time-consuming project with Shreya Narasimhan. Despite the demanding nature of our shared project, her sharp intellect and cheerfulness held up this project and helped make the work a joy. In my early years in the Polley lab, I had the pleasure of working with Blaise Robert who helped me in setting up the imaging microscope and collecting the first imaging of the auditory cortex in awake head-fixed mice. His technical expertise and dry humor was indispensable in pushing that technique development forward. After having set up the rig for head-fixed imaging, the projects started flowing in. There is no way I could have completed all the work presented in chapters 5 and 6 by myself, in a timely fashion, without Sandra Romero. A co-first author on the multi-scale mapping paper, Sandra provided expert and effective analysis and was key for the completion of the project while being an enjoyable colleague to work with. Although we didn’t share the lead on any projects, I appreciate the conversations and support from desk mates Kameron Clayton and Ross Williamson who both helped to widen my perspectives of the auditory system and neuroscience in general. Stephen McInturff and Vivek Kanumuri were ideal members to continue work in the ABI lab after I joined the Polley lab, and I appreciate their graciousness during this transition. Today I continue to consider them among my favorite colleagues.
On a personal level, the last seven years were also a transformation for me. For that I have many groups to thank. I’d like to thank the members of the first SHBT class to matriculate at Harvard, Jennifer Zuk, Rachel Romeo, Brian Buchel, Kevin Woods, and Andrew Ayoob, for helping make this PhD enjoyable, especially during the early years. I consider each one of them as wonderful people and close friends. I’d like to acknowledge my colleagues and faculty in the SHBT program for their friendship and support. One of my vices during my studies was to play soccer, I have many good memories of playing with teammates and games in the intramural and in the Boston Graduate Soccer League. One of my teammates Lee Liu, a brilliant scientist in his own right, was one of those teammates who became close friends and helped balance my life as a student with many nights with soccer, both planned and spontaneous.

At last but not at least, I am indebted to my family. I thank my father, who himself showed me the value in hard work, persistence, and being a wonderful listener. I thank my mother, who provided tireless support emotionally and physically, and who encouraged me to push the boundaries whenever and wherever possible. My brothers, having spent our years together growing up, and time that has undoubtedly shaped me in a positive way. Eric, who today I spend much time musing about philosophy, provocative perspectives on the yin and yang of life, and the impact of relationships while enjoying each other’s company. And Jesse, a brilliant mind of his own, who I have wonderful memories growing up together and despite his proximity far away from Boston continues to inspire me. And I’d like to acknowledge Shreya Narasimhan who not only became my colleague but also my partner, for sharing a wonderful friendship though both science and life, enabling me to become more vulnerable and to grow.
Chapter 2-5 reflect work that is published, and Chapter 6 reflects work that are planned to be submitted for publication. The citations for the respective published work are as follows, in the order they appear as chapters in this dissertation:


Hight, A E., McGill, M., Clayton, K., Parthasarathy, A., Hancock, K., Polley, D (2019). Cortical map dynamics associated with perceptual hypersensitivity following acoustic trauma. To be submitted.

* denotes equal contribution
LIST OF TABLES AND FIGURES

Tables

TABLE 2.1: P-VALUES OF THE MULTIPLE LINEAR REGRESSION TESTS ................................................. 23

Figures

FIGURE 2.1: STIMULATING ARRAY AND RECORDING SETUP .................................................................. 12
FIGURE 2.2: WAVEFORMS AND SPATIAL VARIATIONS ..................................................................... 15
FIGURE 2.3: ACOUSTIC IC RESPONSE MAPS IN 7 INDIVIDUAL RATS SHOW RESTRICTED RESPONSE AREAS WITH CLEAR TONOTOPIC SHIFTS .................................................................................. 19
FIGURE 2.4: ELECTRICAL IC RESPONSE MAPS IN 7 RATS SHOW VARIABLE REGIONS WITH LESS TONOTOPIC SHIFT .................................................................................................................. 22
FIGURE 2.5: SUBTRACTION OF COMMON ACTIVITY AND LONG-LATENCY RESPONSES IN AN INDIVIDUAL RAT (RAT #5). ........................................................................................................... 26
FIGURE 2.6: SUBTRACTION OF COMMON ACTIVITY RESULTS IN LARGER SHIFTS OF CFS AND WIDTHS OF ACTIVATION: ......................................................................................................................... 29
FIGURE 2.7: COMBINED SUBTRACTION OF COMMON ACTIVITY AND LATE PHASE OF RESPONSE SHOWS THE LARGEST SHIFTS OF CFS ........................................................................................................... 33
FIGURE 2.8: RESPONSE MEASURES FOR INTER-ELECTRODE DISTANCES OF 400 VS. 800 µM ............ 35
FIGURE 2.9: SIMILARITIES OF RESPONSES TO MONOPOLAR AND BIPOLAR STIMULATION IN AN INDIVIDUAL RAT (RAT #11). .................................................................................................................... 38
FIGURE 2.10: SIMILARITIES OF RESPONSES TO MONOPOLAR AND BIPOLAR STIMULATION IN ALL RATS. ................................................................................................................................................. 39
FIGURE 3.1: SURGICAL APPROACH TO THE DCN FOR OPSIN INJECTION AND IC FOR NEURAL RECORDING WITH A MULTICHANNEL ELECTRODE PROBE ....................................................................... 44
FIGURE 3.2: CHRONOS EXPRESSION IN THE COCHLEAR NUCLEUS .................................................. 49
FIGURE 3.3: ICC NEURAL ACTIVITY EVOKED BY LIGHT PULSES (CHRONOS INJECTED MOUSE). ....... 51
FIGURE 3.4: SYNCHRONIZATION INDEX AND DRIVEN RATE AS A FUNCTION OF PULSE RATE ......... 52
FIGURE 3.5: NEURAL ACTIVITY AS A FUNCTION OF STIMULUS INTENSITY AND RECORDING POSITION

FIGURE 4.1: CHRONOS-MEDIATED UNIT RESPONSES ARE MORE ROBUST THAN CHR2 OR NATURAL ACOUSTIC STIMULATION.

FIGURE 4.2: CHRONOS SUPPORTS A SUPERIOR NEURAL CODE FOR STIMULATION RATE

FIGURE 4.3: CHRONOS-MEDIATED SPIKING PROVIDES A SUPERIOR BASIS FOR PULSE RATE DETECTION IN THE MIDBRAIN.

FIGURE 4.4: BEHAVIORAL DETECTION OF MIDBRAIN PHOTOSTIMULATION IS SIMILAR ACROSS PULSE RATES AND OPSIN TYPES.

FIGURE 4.5: AUDITORY CORTEX ENCODING OF MIDBRAIN PHOTOSTIMULATION MATCHES BEHAVIOR DETECTION, NOT ICC NEURAL DETECTION.

FIGURE 5.1: A TRANSGENIC GCAMP6S REPORTER MOUSE THAT RETAINS GOOD HEARING INTO ADULTHOOD.

FIGURE 5.2: ANATOMICAL LANDMARKS FOR TONOTOPICALLY ORGANIZED FIELDS IN THE MOUSE AUDITORY CORTEX.

FIGURE 5.3: DATA-DRIVEN PARCELLATION OF AUDITORY CORTICAL FIELDS.

FIGURE 5.4: MAPS AND MODULES IN THE AUDITORY CORTEX.

FIGURE 5.5: TONOTOPIC MAP STABILITY OVER TIME.

FIGURE 5.6: SPATIAL ALIGNMENT OF MESOSCALE AND CELLULAR FREQUENCY TUNING.

FIGURE 5.7: LOCAL AND GLOBAL ORGANIZATION OF TONOTOPY.

FIGURE 5.8: MAPPING THE CORE AND HIGHER-ORDER FIELDS OF THE MOUSE AUDITORY CORTEX.

FIGURE 6.1: HIGH-FREQUENCY NOISE EXPOSURE INDUCES HIGH-FREQUENCY PTS CAUSED BY LOSS OF OHC FUNCTION.

FIGURE 6.2: HIGH-FREQUENCY HEARING LOSS CONFIRMED BY OPERANT BEHAVIOR VIA ADAPTIVE TRACKING WITH FLANKERS PARADIGM.

FIGURE 6.3: BEHAVIORAL EVIDENCE OF ACOUSTIC TRAUMA INDUCED HYPERSENSITIVITY AT SPARED FREQUENCY.
FIGURE 6.4: CALCIUM IMAGING IN AWAKE HEAD-FIXED MICE PROVIDE MAGNITUDE OF TONE-EVOKED ACTIVITY AND TOPOGRAPHICAL ORGANIZATION OF ACTX..................................................152

FIGURE 6.5: HIGH-FREQUENCY ACOUSTIC TRAUMA INDUCES RAPID CORTICAL RE-ORGANIZATION OF SOUND FREQUENCY........................................................................................................155

FIGURE 6.6: CORTICAL REORGANIZATION IN PTS ZONE IS CAUSED BY ACTIVE ENHANCEMENT OF FREQUENCIES AT EDGE OF LESION.................................................................................................157
CHAPTER 1 INTRODUCTION

Disabling hearing loss affects approximately 466 million people worldwide, resulting in impaired speech comprehension and associated hearing disorders such as tinnitus and hyperacusis. One of the most successful interventions for those with complete hearing loss is the modern multi-channel auditory neuroprosthesis. Auditory neuromodulation bypass sensory cells or primary neurons in the inner ear to directly stimulate early stages of the auditory pathway. These devices were first approved for clinical use by the United States Food and Drug Administration (FDA) for adults in 1985 and later for children in 1990. Since then over 320,000 patients have been implanted, marking them as one of the most successful sensory restoration devices to date. These devices reliably restore sound awareness but have a mixed record of success in restoring unaided speech cues that form the basis of oral inter-personal communication. This dissertation uses a mouse model to explore new approaches for stimulating the central auditory pathway and new approaches to understand how auditory regions of the brain reorganize after hearing loss.

Chapters 2-4 explore efforts to identify neural bottlenecks in providing usable auditory cues through electrical stimulation of the brainstem and describe efforts to circumvent these bottlenecks through optogenetic stimulation. For these studies, we focused on the auditory brainstem implant (ABI), a device that bypasses a damaged peripheral auditory system to activate second order neurons located at the cochlear nucleus. The ABI is obstinately similar to the cochlear implant (CI), a device that is implanted into the inner ear and stimulates the primary afferent neurons. The two devices are similar in that the speech processor worn externally on top of the ear, and the internal component that is implanted underneath the skin and within auditory
structures are exactly the same except for the electrode array design. Whereas the CI is a long, narrow, and cylindrical with electrode contacts along the length that slide along the cochlear spiral, the ABI electrode array is flat and rectangular with electrode contacts positioned along its surface. These ABI electrode contacts are placed proximal to second order neurons in the cochlear nucleus (CN) provides patients with general sound awareness but limited speech cues, suggesting limitations in spectral and temporal resolution. In Chapter 2, we designed a micro-electrode array to test whether topographical cues are provided by electrical stimulation delivered to the dorsal cochlear nucleus (DCN). The general summary of findings is that most electrodes evoked low threshold and strong activation of the auditory pathway, but further analysis of the electrically evoked activity in upstream neurons of the auditory midbrain demonstrate limited topographical cues, suggesting a severe bottleneck with electrical stimulation of the CN. Whereas electrical microstimulation indiscriminately activates neurons in central pathways, employing optical stimulation combined with light-activated ion channels that genetically target specific neuronal types (i.e. optogenetics) provide real hope for overcoming limitations of electrical stimulation. Before targeted delivery of channelrhodopsins can be achieved in the central nervous system of human patients a host of technological, ethical, and regulatory hurdles must be overcome. The chief concern regarding the performance of optogenetics in neuroprostheses, however, is its inherent sluggishness in responsivity to rapid light pulses. In Chapters 3 and 4, we were inspired to examine a recently identified channelrhodopsin, nicknamed ‘chronos,’ with fastest known temporal kinetics of any channelrhodopsin, either naturally occurring or engineered. The general summary of these studies is that we found that chronos supports non-adapting, high-fidelity encoding of rapid light pulses within stimulation rates used in ABIs. In Chapter 4, we went one step further and asked
whether different pulse rates would translate to differences perception. Using the auditory midbrain as a testbed due to its accessible anatomy, using a chronically implanted fiber to deliver optical pulses we found a clear neural code for stimulation rate in the auditory midbrain but found no differences in behavioral responses. This set of studies in chapters 2-4 establish potential bottlenecks in the neural code of both topographical and temporal stimulation when examined at cortical areas higher than the location of stimulation and that optical stimulation continues to be a promising avenue for breakthrough improvements in auditory neuroprostheses. One of the most interesting findings of chapter 4, however, was the dramatic transformation from a sustained rate and timing code to a transient and fully adapting rate code in individual neurons of the auditory cortex. These findings suggested future studies in the cortical encoding of sound and how this may relate to perception.

Chapters 5 and 6 describe and leverage a new approach to study the organization of downstream processing area, the auditory cortex, in mice with normal hearing and in mice with sudden traumatic hearing loss. We designed and implemented a preparation for chronic wide-field epifluorescence imaging of genetically encoded calcium indicators in excitatory neurons of the auditory cortex in awake head-fixed mice. Using this preparation, Chapter 5 implemented a detailed analysis of the organization of sound frequency across scales, cortical areas, and days. In summary, these analyses provided evidence for alignment of mesoscale and cellular-scale topography, evidence for a medium degree of tuning heterogeneity among nearby cortical neurons, and identification of two areas as potential candidates for identification as secondary auditory fields. We use findings of chapter 5 in chapter 6 to guide our investigation into compensatory plasticity following sudden high-frequency acoustic trauma and associate this plasticity with a behavioral hypersensitivity at frequencies unaffected by the hearing loss.
Chapters 5 and 6 address questions about both, the organization of sound coding in the auditory cortex, reveal differences in this sound coding across different primary areas, and show a tight association between cortical codes of sound representation and perception.
CHAPTER 2

AUDITORY BRAINSTEM STIMULATION WITH A CONFORMABLE MICROFABRICATED ARRAY ELICITS RESPONSES WITH TONOTOPICALLY ORGANIZED COMPONENTS


Abstract

Auditory brainstem implants (ABIs) restore hearing to deaf individuals not eligible for cochlear implants. Speech comprehension in ABI users is generally poor compared to that of cochlear implant users, and side effects are common. The poor performance may result from activating broad areas and multiple neuronal populations of the cochlear nucleus, however detailed studies of the responses to surface stimulation of the cochlear nucleus are lacking. A conformable electrode array was microfabricated to fit on the rat’s dorsal cochlear nucleus (DCN). It hosts 20 small electrodes (each 100 µm diam.). The array was tested by recording evoked potentials and neural activity along the tonotopic axis of the inferior colliculus (IC). Almost all bipolar electrode pairs elicited responses, in some cases with an even, or relatively constant, pattern of thresholds and supra-threshold measures along the long axis of the array. This pattern suggests that conformable arrays can provide relatively constant excitation along the surface of the DCN and thus might decrease the ABI side effects caused by spread of high current to adjacent structures. We also examined tonotopic patterns of the IC responses. Compared to sound-
evoked responses, electrically-evoked response mappings had less tonotopic organization and were broader in width. They became more tonotopic when the evoked activity common to all electrodes and the late phase of response were subtracted out, perhaps because the remaining activity is from tonotopically organized principal cells of the DCN. Responses became less tonotopic when inter-electrode distance was increased from 400 µm to 800 µm but were relatively unaffected by changing to monopolar stimulation. The results illustrate the challenges of using a surface array to present tonotopic cues and improve speech comprehension in humans who use the ABI.

**Keywords.** Auditory brainstem implant; ABI; neural stimulation; electrical stimulation; microelectrode; auditory neuroprosthesis

**Highlights**

- A microfabricated conformable electrode array evoked responses when placed on the DCN surface in rats.
- The regional pattern of eABR and neural IC responses showed little variation along the long axis of the array, suggesting the array provided uniform electrical stimulation along the lateromedial axis of the DCN.
- Compared to sound-evoked responses, electrically-evoked responses had less tonotopic organization and were broader.
- The tonotopy became greater after subtracting the evoked activity common to all electrodes and the late phase of response, perhaps isolating activity from tonotopically organized principal cells of the DCN.
Introduction

The auditory brainstem implant (ABI) is a neuroprosthesis that provides sound sensations by stimulating neurons of the cochlear nucleus (CN) in deaf patients who are not candidates for a cochlear implant (CI). The ABI has been approved by the Food and Drug Administration for adults with Neurofibromatosis Type 2 (NF2), a devastating genetic syndrome associated with bilateral vestibular schwannomas. The growth or removal of these tumors results in deafness from injury to the cochlear nerves. Also candidates for the ABI are children with hypoplasia of the cochlea or cochlear nerve, and adults and children with scarring of the cochlea following trauma, otosclerosis, or meningitis (Colletti et al., 2009a; Noij et al., 2015; Puram et al., 2016). Altogether, over 1,500 patients have been implanted since 1979, when the first patient, a woman with NF2, received a device with two ball electrodes placed on the CN (House et al., 2001).

Only a small group of ABI users achieve the ultimate goal of an auditory prosthesis, open set sentence comprehension, which is the ability to understand speech without visual clues. The majority of ABI users only obtain general sound awareness that does aid in lip reading (Schwartz et al., 2008; Colletti et al., 2009b; Vincent, 2012). In contrast, most CI users attain much better performance and achieve some degree of open set speech comprehension. Psychophysical experiments show that ABI and CI users perform comparably in detecting temporal cues (i.e. amplitude modulated signals), but ABI users perform much poorer at discriminating between frequencies (Shannon and Otto, 1990) and can have trouble ranking the pitches of their electrodes (Otto et al., 2002). Although contemporary ABI devices consist of an array of 12 to 21 electrodes placed on the surface of the CN, accurate frequency encoding may not be provided because of broad, overlapping regions of excitation of subsurface neurons. In addition, multiple
subtypes of neurons in the CN are likely to be stimulated by the ABI in a mixture of tonotopic and non-tonotopic patterns. Such simultaneous activation likely contributes to minimal speech comprehension in ABI users. Other factors limiting ABI performance could be damage to the CN caused by a tumor and/or its removal (Colletti et al., 2005) and non-optimal placement of the ABI electrode array (Barber et al., 2017).

The surface stimulation of the CN used in the human ABI has not previously been studied in detail in animal models. One study using two surface electrodes (McCreery et al., 2010) showed that responses are broader than those evoked by microstimulation within the CN using penetrating electrodes (McCreery et al., 2010; Shivdasani et al., 2008; Takahashi et al., 2005). Penetrating electrodes were used in a few human subjects but have been discontinued because of side effects (Otto et al., 2008). Surface electrodes almost certainly evoke responses in multiple types of CN neurons. One type that is tonotopically organized in the dorsal cochlear nucleus (DCN) is the pyramidal/fusiform neuron; these principal neurons are arranged from low to high characteristic frequency (CF) along the long axis of the DCN (Kaltenbach et al., 1991; Oertel et al., 2004; Ryan et al., 1988; Spirou et al., 1993). Thus, ventrolateral electrodes are likely to activate mostly low-CF principal neurons and dorsomedial electrodes mostly high-CF principal neurons. However, this may not be the case for DCN interneurons, for fibers, and for the more remotely located neurons of the ventral subdivision (VCN). The present study used a strategy to tease out tonotopic patterns by subtracting activity that is non-tonotopic because it was “common”, or elicited in response to most or all electrode pairs. A second strategy was to subtract the late phase of response, because the early phase of response is likely to be from principal cells that have large-diameter, fast-conducting axons (Adams, 1979; Baizer et al., 2012; Schofield et al., 2005).
A further strategy to improve performance of surface stimulation was to use a flexible electrode array (Guex et al., 2015). This type of array can conform to the small radius of curvature of the rat DCN (radius = 1.65 mm) without a change in electrode impedance. The microelectrodes within the array are embedded in a few micron-thick plastic film. Microfabrication using thin-film micromachining techniques and polymers allows for the miniaturization of the ABI electrode array so that it can fit the small DCN in animal models without compromising electrode density. Indeed, the human clinical ABI made by Cochlear Corp. hosts 21 electrodes on an array that is 3 mm by 8 mm, while that used here carries 20 microelectrodes on an array that is 1.1 mm by 1.9 mm and fits the exposed DCN in the rat. Furthermore, its small electrodes (100 µm in diameter) are also smaller compared to the 550-700 µm diameter electrodes in the human ABI (Vincent, 2012). Finally, we also compared bipolar vs. monopolar stimulation along the CN’s tonotopic axis. These tests are motivated by the idea that bipolar stimulation produces a more focused excitation volume of the neural tissue, as observed in the cochlea (Middlebrooks et al., 2007; Snyder et al., 2008; Zhu et al., 2012) and in deep brain stimulation (Kuncel et al., 2004; O’Suilleabhain et al., 2003). Although bipolar stimulation is not always associated with functional benefits in these studies, it can potentially reduce the lateral spread of stimulation current (Bonham et al., 2008) and might thus reduce side effects.

Methods

Fabrication of the stimulation array

Electrode arrays were fabricated using standard microfabrication processes (Guex et al., 2015). A sacrificial layer of Ti (10 nm)/Al(100 nm) was first deposited by evaporation on a
silicon wafer. A first layer of 4 µm thick polyimide (PI2611, HD Microsystems GmbH, Germany) was then spin-coated and cured (soft bake, 5 min at 120°C followed by hard bake, 2 hours at 300°C in a N₂ oven). Next, the interconnects layer (Ti/Pt/Ti, 75/350/75 nm) was sputtered after O₂ plasma surface activation and patterned by photolithography and IBE (ion beam etching). A second layer of PI (4 µm thick) was subsequently spin-coated and cured, then patterned by photolithography and RIE (reactive ion etching), defining both the electrode active sites and the array’s external shape. The electrode arrays were released from the wafer by anodic dissolution of the Al layer (1 V bias, in saturated NaCl solution). Each array contains 20 electrodes patterned in a hexagonal configuration, with 4 rows, each of 5 electrodes (Fig. 2.1C). The total surface of the array is 1.1 mm by 1.9 mm, which was tailored to be the size of the exposed DCN in rat after cerebellar aspiration as observed in this study. The 100 µm diameter electrodes, patterned on fingerlike stripes of PI, conform to the surface of the DCN (Fig. 2.1A-C). The array interconnects are routed to the side at caudal aspect of the DCN and are interfaced to a rigid cable and connector.

**PEDOT coating of the electrodes**

PEDOT: PSS was electropolymerized from a solution of EDOT (0.1%w/v) and PSS (0.2%w/v) (both purchased from Sigma Aldrich), as described (Guex et al., 2015). Electropolymerization was performed galvanostatically with a deposition charge of 75 mC/cm². The impedance modulus at 1 kHz of the coated microelectrodes was decreased by more than one order of magnitude compared to uncoated microelectrodes.

**Surgery**
All experimental procedures were performed in accordance with the National Institute of Health guidelines for the care and use of laboratory animals as well as approved animal care and use protocols at the Massachusetts Eye & Ear Infirmary, Boston, MA. *In vivo* tests were performed on male Sprague-Dawley rats (350-500 g) within a sound-attenuating chamber. Animals were anesthetized with Ketamine (100 mg/kg) and Xylazine (20 mg/kg) for the duration of the experiment. The head was fixed, a left occipital craniotomy was performed, the dura mater was opened, and a small portion of the ipsilateral cerebellum was aspirated to expose the left DCN (Fig. 2.1A). A second craniotomy was performed over the right temporal-parietal suture, to access the dorsal surface of the right IC.

**Placement of the stimulation array**

The array was placed on the exposed surface of the left DCN (Fig. 2.1A). eABRs (Fig. 2.2A) were recorded differentially using two subcutaneous stainless steel needle electrodes, one placed on the vertex, the other behind the ipsilateral pinna, with a ground electrode placed on the back of the animal. Recordings were averaged over 512 stimulus presentations and filtered with an analog bandpass filter (30 Hz to 3 kHz) amplified by 60 dB before A/D conversion with a sampling frequency of 25 kHz (Ithaco Model 1201, DL Instruments, Ithaca NY). Post-experiment histology confirmed accurate placement of the ABI electrode on DCN (Fig. 2.1B) by inserting pins (0.1 mm diameter Minutien) at the corners of the array and subsequent recovery of the pin-holes in Nissl-stained frozen sections.
Figure 2.1: Stimulating array and recording setup

A) Photograph of the miniaturized, conformable stimulating electrode array in position on the left DCN in a rat. B) Overlaid electrode array schematics from 5 rats, pictured on a color CN reconstruction (from Verma et al., 2014). Schematics were positioned according to post-experiment histology (see Methods). The colored arrow indicates the orientation of the lateromedial axis along the roughly lateroventral-to-mediodorsal dimension of the DCN. C) Photograph of electrode array before positioning, showing the electrode pair spacings and orientations that were tested. D) Schematic showing the recording probe in the IC contralateral to the stimulated side. In the IC, the tonotopic axis runs from low characteristic frequencies (CFs) dorsolateral (dark colors) to high CFs ventromedial (light colors). Left-ear sounds presented at the beginning of each experiment confirmed placement along the tonotopic axis of the IC and determined the CF of the individual recording sites. Subsequently, electrical stimulation of electrode pairs in the DCN evoked IC responses as well as eABRs (recorded by cutaneous electrodes that are not shown).

Recordings of evoked multi-unit activity in the IC
A single-shank, 16-channel penetrating recording probe (A1x16-5 mm -150-177, Neuronexus, Ann Arbor, MI) was inserted into the right IC, the side contralateral to the stimulated DCN (Fig. 2.1D). Multi-unit activity was measured simultaneously from the 16 recording sites, first in response to acoustic stimulation (left ear tone pips, 20 ms duration; 1-46.5 kHz with 2 steps per octave; 0-80 dB SPL). Measured signals were sampled at 25 kHz and the average signal of all sites was used as a global reference and subtracted before further analysis. Responses were displayed as IC response maps (Fig. 2.3) in response to each of the acoustic frequencies. Stimulation with high acoustic frequencies induced activity in the low-numbered recording sites (located ventromedially in the IC); as frequency was decreased, the activity progressively shifted to higher-numbered sites (more dorsolaterally; Figs. 2.1D and 2.3A) and a monotonic progression of response was evidence of the probe’s location in the central nucleus of the IC. The logarithmic fit was used to obtain the characteristic frequency (CF) associated with each recording site (Fig. 2.3B, curves). Only the sites for which acoustically-evoked activity was recorded were considered for analysis of the electrically-generated data.

**Electrical stimulation protocol**

Electrical stimulation was induced with custom-made isolated current sources. Stimuli were current-controlled but expressed in voltage levels due to the inadvertent presence of a 2 kΩ resistor in parallel with the output during the experiments. Thus, although the exact stimulation current levels are unknown, they are estimated from combined electrode/tissue impedances. At the end of the experiment, we measured the electrode-tissue impedance systematically across the array with a PalmSens3 impedance analyzer and found average values of 395 kΩ at 1 kHz and 74 kΩ at 10 kHz (n = 17 electrodes). In saline solution, the typical electrode impedance was
measured to be 2.66 kΩ at 1 kHz (Guex et al., 2015). Thus, a stimulation at 4 V (the maximal used) corresponds to a maximum current level of 0.01 mA flowing into the DCN.

Trains of charge-balanced biphasic symmetric pulses at 0.2 ms/phase were delivered to the CN at 28 pulses/s and alternating in polarity. Minimal differences were observed between the two different polarities so the average of the two polarities was used in the analysis presented here. Stimulation level ranged from 0 to 4V in steps of 0.5 V. Baseline activity was recorded during the 500 ms off period after each 500 ms pulse train. Bipolar stimulation was applied between two electrode sites of the array, whereas monopolar stimulation was applied between one electrode in the array and a reference electrode (stainless steel needle placed in the neck muscles of the animal). Due to the hexagonal structure of the electrode sites on the array, three orientations were characterized, and defined by their angle with respect to the lateromedial axis (-60°, 0, 60°). Two inter-electrode distances were tested, 400 μm and 800 μm. Monopolar and bipolar stimulation runs were performed in a pseudo-random order.
Figure 2.2: Waveforms and spatial variations

A) Example electrically-generated auditory brainstem response (eABR) showing 4 waveform peaks. The stimulation pulse started at time 1 ms and in this trace the stimulation artifact was very small. B) eABR thresholds (left) and suprathreshold amplitudes (right) plotted as color-coded matrices, replicating the view of the array on the DCN (Fig. 

...
2.1A-C). The matrix x and y axes are the lateromedial and rostrocaudal dimensions, respectively (see arrows). Data are for individual Rat #5. Suprathreshold amplitudes are eABR normalized RMS energy (of the first 10 ms of the signals 1 V above threshold). NF represents a non-functional electrode on this array.  

C) Example multi-unit waveform from one IC recording site after filtering. The trigger level for spike detection is shown by the green horizontal line. The first 1 ms of the trace, which contains the stimulation artifact, was deleted (spike latencies were longer than 1 ms). D) Color-coded matrices as in B but here for IC spike count threshold and a suprathreshold measure (normalized IC spike count at 1V above threshold). Thin dashed lines represent electrode pairs producing no significant response. E) Comparison of lateromedial spatial variations of eABR and IC thresholds (left) and suprathreshold spike count (right), for the individual rat (panels B,D) and for the average of all rats. “Coordinate” on the x axes refers to distance from center of the array (0) to the midpoint of the bipolar electrode pair (see matrices above) with negative values lying laterally and positive values medially (arrow). F) Rostrocaudal spatial variations of eABR and IC threshold (left) and suprathreshold spike count (right). The x-axis here (rostrocaudal dimension) was the y-axis of the matrices (panels B,D). Negative coordinate values lie rostrally and positive lie caudally (arrow).

**Multi-unit recordings processing and analysis**

Processing of the data was performed with Matlab software (The Mathworks Inc., Natick, MA). The first 1 ms following each stimulus pulse onset was discarded to eliminate the stimulation artifact. A forward and reverse bandpass Butterworth filter of order 5 was applied, with cut-off frequencies of 500 Hz and 3 kHz. Spikes were detected by non-negative crossings of a threshold (Fig. 2.2C) set at three times the standard deviation of spontaneous activity. The evoked spike rate was determined after subtraction of the baseline spike rate. Averages of 30 repetitions of 500 ms pulse trains were obtained for each condition.

Spike rates were averaged across trial duration and stimulus repetitions. Post-stimulus time histograms (PSTHs) were obtained for each condition with a resolution of 0.1 ms.
Significance of activation for each recording site and stimulus intensity combination was determined by a student’s t-test between spike rate during the stimulus on and stimulus off periods. Spike rate of non-significant conditions (p<0.001) were set to zero to decrease the noise of the calculations.

The best threshold for each IC response map was determined as the lowest electrical stimulation level generating significant activity on at least one of the recording sites. For each level \( l \) above best threshold, the centroid of activity was calculated with the following formula (\( k: \) IC recording sites; \( MAP: \) 2-dimensional IC response map):

\[
C(l) = \frac{\sum_k k \cdot MAP(k, l)}{\sum_k MAP(k, l)}
\]

The width of the IC response map was defined for each level above best threshold as the double of the standard deviation of the centroid (Escabi et al., 2007), with IC recording site instead of octave frequency):

\[
W(l) = 2 \cdot \sqrt{\frac{\sum_k (k - C(l))^2 \cdot MAP(k, l)}{\sum_k MAP(k, l)}}
\]

The centroid for electrical stimulation was defined at a level of 0.5 mA above the best threshold (\( C_e(T + 0.5 \, mA) \)) for each IC response map. The centroid for acoustic stimulation was defined at 10 dB SPL above threshold (\( C_a(T + 10 \, dB \, SPL) \)) for each acoustic frequency. Rate-level curves were obtained using the average spike rates at the centroid ±1/2 width. Spike counts were obtained at each level by adding the evoked spikes over all IC recording sites. Rate-level curves and spike counts were normalized for each rat to the maximum of the rate level curves, respectively the maximum spike count across all stimulation conditions.

The common activity was calculated by averaging IC response maps with significant activity evoked by all electrode pairs across the DCN. This average map was calculated
separately for the different inter-electrode distances. The early vs. late phase of response were
distinguished qualitatively by PST histograms.
Figure 2.3: Acoustic IC response maps in 7 individual rats show restricted response areas with clear tonotopic shifts

A) Response patterns in the IC following acoustic stimulation at 1, 4, 8, 16 and 32 kHz shown for each of the seven rats. These response maps have colors that correspond to firing rates (color scales) as a function of stimulation voltage and plotted for each of the 16 IC recording sites. A shift in activity along the recording probe is observed, from low site numbers at high frequency to high site numbers at low frequency. In top, left panel, red arrow indicates centroid of activity. B) The mapping of acoustic frequency vs. centroid is the tonotopic mapping. Data points show the IC centroids corresponding to each acoustic frequency tested. Lines show logarithmic fit to the data, which determines the characteristic frequency (CF) for each IC recording site.

Statistical analyses

Results are expressed as mean ± standard error of the mean. Statistical analyses were performed on the data and significance was indicated by a p-value lower than 0.05. Pearson’s correlation, multiple regression analysis, analysis of covariance (ANCOVA) and ANOVA were used, depending on the data.

Results

Responses to DCN electric stimulation

With the array placed on the surface of the DCN, all electrodes were observed to be in close contact with the neural tissue (Fig. 2.1A). Electrical stimulation using pairs of electrodes evoked eABRs (Fig. 2.2A). The eABR waveform was variable, even for different electrode pairs within the same animal, and it consisted of 2 – 4 positive-going and negative-going peaks in succession in the 5 ms after the electrical pulse (Fig. 2.2A). The eABRs showed little change in threshold or suprathreshold amplitude with change in electrode location along the lateromedial (long) dimension of the array (as shown by matrices for one rat in Fig. 2.2B, and by graphs for
all data in red in Fig. 2.2E). In contrast, along the rostrocaudal dimension, thresholds decreased and suprathreshold amplitudes increased caudally (Fig. 2.2F).

Electrical stimulation also evoked multiunit neural activity in the IC (Fig. 2.2C), with rates significantly greater than spontaneous activity in the majority of runs (372 out of 398, 93.5% across all animals). Like the eABRs, there was little change in multi-unit activity evoked by electrodes positioned along the lateromedial dimension (Fig. 2.2D, E). In contrast, rostrocaudally, thresholds increased somewhat and normalized spike rate decreased caudally (Figs. 2.2F and Table 1), which is a pattern opposite to that for the eABR. Electrode pairs with the highest IC thresholds or no responses were almost always located on the rostral edge, which is shown by the darker colors toward to the tops of the matrices in Figure 2.2D. Post-experiment histology suggests that these rostral-edge electrodes were at or just beyond the edge of the DCN (Fig. 2.1B).

**Tonotopic patterns in raw responses**

*Patterns in responses evoked by sound.* Each rat showed responses to individual sound frequencies over a narrow range of recording sites in the IC (Fig. 2.3A). Response maps were V-shaped. For each sound frequency, we calculated a centroid of activity that was usually at the tip of the V (Fig. 2.3A, color IC response map at top left). Plots of the centroids (Fig. 2.3B) showed orderly mappings that were similar across animals with an average of 1.92 recording sites per octave of frequency. The logarithmic fit to the data (Fig. 2.3B, curves), the tonotopic mapping, was used to assign a characteristic frequency (CF) to each recording site in the IC (see Fig. 2.4B, x axis), and these CFs were used to determine the mapping of response to electrical stimulation (Fig. 2.4C-E).
Patterns in responses evoked by electrical stimulation. Compared to the patterns in response to sound, those in response to electrical stimulation had more variability. The response areas were broader and some areas were “split” between two groups of IC recording sites (Fig. 2.4A, Rat #4). The centroid of response to electrical stimulation (Fig. 2.4B) varied with stimulus pair position along both the lateromedial and rostrocaudal dimensions of the DCN (Figs. 2.4C, D). The axis that maximizes the change in CF, the tonotopic axis (Fig. 2.4E), is a combination of the lateromedial and rostrocaudal array dimensions, consistent with the spatial mapping of sound-evoked responses cats (Spirou et al., 1993) and mice (Muniak et al., 2013). The mappings across all stimulating electrodes, though, showed ranges (avg. 1.03 octaves, 17.23 to 35.3 kHz) smaller than for sound stimulation (5.54 octaves, 1 to 46.5 kHz). They could be horizontal (e.g., Rat 1 in Fig. 2.4C), which indicates no change in CF with change in electrode pair position lateromedially. Overall, response CF was positively correlated with the electrode pair position in only 4/7 rats lateromedially but in 6/7 rats rostrocaudally (Fig. 2.4C, D, Table 1).
Figure 2.4: Electrical IC response maps in 7 rats show variable regions with less tonotopic shift

A) Top: Five bipolar pairs of stimulating electrodes (red, spacing 400 µm) chosen because they span the lateromedial and rostrocaudal distances along the DCN. Bottom: Response maps (as in Fig. 2.3) to electrical stimulation applied by the five bipolar pairs in seven rats. B) Example response map indicating the features that were extracted for analysis (see Methods), with centroid and width computed at 1V above best threshold. The scale at the bottom shows the tonotopic mapping of IC recording sites to CF, which was determined for each animal using acoustic stimulation (see Fig. 2.3 and Methods). C, D) CF as a function of DCN coordinate for all bipolar pairs
Figure 2.4 (Continued)

(averages of all orientations at the particular distance point) across the seven rats. Response CF is positively correlated with the lateromedial electrode pair position in rats 2, 3, 4 and 5 and with rostrocaudal position in rats 1, 2, 3, 4, 5 and 6 (see Table 1). E) CF as a function of distance along the tonotopic axis, which was determined by performing a multiple linear regression of CF vs. distance along the lateromedial and rostrocaudal axes. Each point corresponds to a bipolar pair.

Table 2.1: P-values of the multiple linear regression tests.

P-values of the multiple linear regression of the best threshold (T), normalized spike count (NSC), CF, and width of activation (W) for lateromedial (lm) and rostrocaudal (rc) position of the stimulation pair. Values are also given for the interaction factor (int).

<table>
<thead>
<tr>
<th></th>
<th>T-lm</th>
<th>T-rc</th>
<th>T-int</th>
<th>NSC-lm</th>
<th>NSC-rc</th>
<th>NSC-int</th>
<th>CF-lm</th>
<th>CF-rc</th>
<th>CF-int</th>
<th>W-lm</th>
<th>W-rc</th>
<th>W-int</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.505</td>
<td>0.382</td>
<td>0.015</td>
<td>1.4E-5</td>
<td>0.034</td>
<td>0.030</td>
<td>0.218</td>
<td>3.9E-3</td>
<td>1E-4</td>
<td>0.102</td>
<td>7.7E-3</td>
<td>5.7E-3</td>
</tr>
<tr>
<td>2</td>
<td>0.543</td>
<td>0.174</td>
<td>0.550</td>
<td>0.010</td>
<td>0.024</td>
<td>5.8E-3</td>
<td>3.1E-3</td>
<td>2.1E-3</td>
<td>0.092</td>
<td>0.060</td>
<td>5.6E-3</td>
<td>0.079</td>
</tr>
<tr>
<td>3</td>
<td>0.078</td>
<td>0.316</td>
<td>0.202</td>
<td>0.021</td>
<td>1.6E-4</td>
<td>0.833</td>
<td>0.029</td>
<td>9E-8</td>
<td>0.752</td>
<td>0.281</td>
<td>1.2E-7</td>
<td>0.060</td>
</tr>
<tr>
<td>4</td>
<td>0.220</td>
<td>0.021</td>
<td>0.478</td>
<td>1.4E-4</td>
<td>3.7E-5</td>
<td>0.065</td>
<td>1.7E-5</td>
<td>0.023</td>
<td>0.073</td>
<td>0.308</td>
<td>0.661</td>
<td>7.5E-3</td>
</tr>
<tr>
<td>5</td>
<td>3.8E-3</td>
<td>3.5E-6</td>
<td>0.750</td>
<td>1.4E-9</td>
<td>2.2E-13</td>
<td>2E-3</td>
<td>8.9E-5</td>
<td>1.4E-8</td>
<td>0.102</td>
<td>5.5E-4</td>
<td>1.2E-8</td>
<td>0.368</td>
</tr>
<tr>
<td>6</td>
<td>2.3E-3</td>
<td>2.6E-3</td>
<td>0.920</td>
<td>5.2E-9</td>
<td>5.6E-22</td>
<td>1.9E-3</td>
<td>0.079</td>
<td>8.8E-8</td>
<td>0.647</td>
<td>0.461</td>
<td>2.3E-6</td>
<td>0.809</td>
</tr>
<tr>
<td>7</td>
<td>0.536</td>
<td>4.1E-3</td>
<td>0.027</td>
<td>4.1E-8</td>
<td>1.4E-7</td>
<td>1.2E-4</td>
<td>0.144</td>
<td>0.155</td>
<td>0.564</td>
<td>3.9E-4</td>
<td>0.030</td>
<td>0.382</td>
</tr>
</tbody>
</table>

Tonotopic patterns after subtraction of common activity and late phase of response

Common activity. Such weak mappings could result from the responses being dominated by non-tonotopic activity that is evoked by all electrode pairs. To investigate this idea, we subtracted the “common activity” from the overall response. The common activity was
calculated as an average IC response map; across rats, these maps had an average CF of $22.4 \pm 1.5$ kHz. The average map for one rat is shown in Figure 2.5A (center IC response map). In this rat, the effect of subtraction on the response areas to two stimulus pairs was to narrow them and to shift their centroids to somewhat higher CFs (Fig. 2.5A left, red numbers on response maps). For all stimulus pairs, matrices (Fig. 2.5A right) show CFs for the raw data and after subtraction. Before subtraction, the range of CFs was 1.39 octaves (12 to 31.6 kHz), whereas after, the range of CFs was 2.67 octaves (8.1 kHz to 51.5 kHz). Combined data (Fig. 2.6A) also showed this effect: a larger range (1.91 octaves after subtraction vs. 0.83 octaves before). Thus, the plot of CF vs. distance along the lateromedial (Fig. 2.6B) and tonotopic axes (Fig. 2.6B inset) both had significantly greater slopes after subtraction (ANCOVA). The direction of the tonotopic mapping, or “CF shift orientation”, was $68^\circ$ for the raw data and was $51^\circ$ after average subtraction.

*Early vs. late phase of response.* Weak mappings from electric stimulation might also result from activation of less tonotopic neurons having long latency. In contrast, short latency is a characteristic of the fast-conducting axons of the tonotopic DCN principal cells. Using latency, the response was separated into early phase and late phase activity using PST histograms (e.g. Fig. 2.5B). The cutoff between early vs. late phase of response was quite variable across electrode pairs (both within a single test subject and between subjects); its value ranged between 4.75 and 9.75 ms. The IC response maps in Figure 2.5B show that the early phase had narrower response areas than the late phase. The early phase also activated IC recording sites over a greater range of CFs. In the case of the rat shown in Figure 2.5B, early-phase CFs had a range of 2.67 octaves (8.1 kHz to 51.4 kHz), whereas for the late phase CFs had a range of only 1.47 octaves (14 to 36.1 kHz), as shown in the matrix plots. Data from all seven rats showed this
effect, with an average range of 1.20 octaves for early vs. 0.89 octaves for late, so that the CF-distance plot for early spikes showed a greater slope (Fig. 2.6E) and was significantly different with lateromedial position although not with tonotopic position (ANCOVA). Perhaps these different effects result from the large change in angle of the CF shift orientation (4° for early phase vs. 76° for the late phase).

*Combined subtractions.* A combined subtraction of the common activity and the late phase of response generated the CF plot with the steepest slope (Fig. 2.7A). This range of CFs (1.96 octaves) was slightly larger than just the average subtraction (1.88 octaves), and clearly larger than for the early phase of response, (1.07 octaves), with both significantly different than the combined (ANCOVA). The average CF shift orientation after combined subtraction was 12° (Fig. 2.7A inset), close to the CF shift orientation obtained with early latency spikes (4°) but very different than raw data (68°) and average subtraction (51°).
Figure 2.5: Subtraction of common activity and long-latency responses in an individual rat (Rat #5).

A) Center: Average IC response map to all bipolar electrode pairs in this animal. Left: two response maps using two bipolar pairs of stimulating electrodes (red, above) that differ in lateromedial position. The raw responses and the responses after subtraction are shown. Subtraction results in narrowing of the response areas and small shifts of CF (red numerals). Matrices to the right: CFs for all electrode pairs using color-coded segments (see scales) between the two electrodes of each pair. The pattern illustrates the CFs across the DCN’s lateromedial (x axis) and rostrocaudal (y axis) dimensions before (above) and after (below) the subtraction of common activity using the average response map. Thin dashed lines represent electrode pairs producing insufficient response to estimate the CF. NF represents a non-functional electrode on this array. B) Center PST: Responses separated into early and late phases according to a latency cutoff. Left: Response maps showing early phase is narrower than late phase for the two electrode pairs at top. Matrices on the right show the CF pattern for all pairs. Thin dashed lines represent electrode pairs producing insufficient response to estimate the CF.

Widths of response areas
In raw, unprocessed responses, widths of IC response maps (Fig. 2.4A) to electrical stimulation were variable and usually broad. They averaged 6.4 ± 0.3 recording sites (3.33 octaves). Widths were significantly correlated to lateromedial position only in rats 5 and 7, with a narrower width on the lateral part of the implant. Width had significant rostrocaudal dependence for 6/7 rats (all except Rat 4, Table 1), with narrower width on the rostral side of the implant. After subtraction of the common activity, response maps became on average significantly narrower (multiple linear regression, p<0.05, Fig. 2.5A), and the width showed more dependence on lateromedial position of electrode pair (Fig. 2.6C). The width was also significantly narrower for early phase compared to the late phase of response (mult. lin. reg, p<0.05, Fig. 2.5B), but both phases showed almost no dependence on lateromedial position (Fig. 2.6F). The width after combined subtraction was narrower than after each individual subtraction (p < 0.05 in both cases, Fig. 2.7B). As in the data after subtraction of common activity, the width was significantly correlated to the lateromedial position (p < 0.05 in both cases), with lower widths at lateral positions. Overall, combining both subtractions produced response maps with a maximal range of CFs, minimal widths, and resulted in a near-lateromedial orientation of the tonotopic axis.

Effects of stimulating electrode distance and pair orientation

So far, data presented have used bipolar stimulation with an inter-electrode spacing of 400 µm. The IC responses to larger inter-electrode distance (800 µm), with the combined subtraction, had fewer rats (4/7 at 800 µm vs. 6/7 at 400 µm) with a significant correlation between CF and lateromedial position. The range of CFs was somewhat smaller (32.9 kHz vs. 39.6 kHz). The only effect of electrode pair orientation was that at 800 µm distance (Fig. 2.8B), the range of CFs was somewhat larger at a -60° orientation. However, effects of angle at either
400 or 800 µm spacing were not significant (ANCOVA). Larger inter-electrode distance also broadened the width of activation, for all orientations combined (ANOVA, p<0.05). For a 400 µm distance, the width was on average 3.17 recording sites (1.65 octaves), whereas for 800 µm the width was on average 3.6 sites (1.88 octaves, Fig. 2.8D). Again, the only effect of electrode-pair orientation was for 800 µm distance, and again, the pairs with a -60° orientation were different (they generated a narrower width than the other orientations, ANOVA, p<0.05).

Variations of threshold and normalized spike count with distance and orientation showed no significant effect (Fig. 2.8C,-E; ANOVA, p>0.05). In the raw data, however, there was a significantly larger spike rate with 800 µm distance (not illustrated).

**Relative similarity of bipolar vs. monopolar stimulation**

Response patterns to monopolar stimulation were generally similar to bipolar stimulation. These experiments used an additional 4 rats and 400 µm pair spacing for direct comparison. Figure 2.9A shows in one rat the similar shapes of raw IC response areas for bipolar stimulation vs. monopolar stimulation (5 bipolar pairs vs. 10 corresponding monopolar electrodes). Similar spatial variations were observed both in raw data (Figs. 2.9B, 10A) and data with combined subtraction (Figs. 2.9C, 10B). There was no significant difference in the CF variation with the lateromedial position of stimulation between monopolar and bipolar data (multiple linear regression, p>0.05 both for the first order polarity effect and for the interaction polarity / position).

The main difference between bipolar and monopolar stimulation was that spike rates were significantly higher for the latter after combined subtraction (Fig. 2.10E; ANOVA, p<0.05). This effect was also observed in the raw data but did not reach significance (p>0.05). No significant
differences were observed in best threshold or width of activation, both before and after subtractions (Fig. 2.9C, D; ANOVA, p>0.05).

**Figure 2.6: Subtraction of common activity results in larger shifts of CFs and widths of activation:**

Early phase responses have larger shifts and narrower widths. A) Color-coded matrices of CFs (left) before (top) and after (bottom) subtraction of average activity. The CF shift orientations (dotted lines) are shown. B, C) Plots of CF and width before and after average subtraction as a function of lateromedial position or tonotopic axis position (inset). After average map subtraction, the response CF of individual maps is positively correlated with the electrode pair position along the lateromedial axis in 5/7 rats (vs. 4/7 before subtraction). Multiple regression analysis demonstrate significant interactions between before vs. after average subtraction and the lateromedial coordinate (p < 0.05). Shaded areas are calculated with the regression of the top and bottom of the standard errors. D, E) Similar data after separation of early and late phases of response. Multiple regression analysis demonstrates significant interactions between early vs. late and the lateromedial coordinate (p<0.05).
Discussion

Present results are the first detailed tests of surface stimulation in an animal model of the human ABI. The unprocessed responses had limited tonotopy and were usually wide. These results offer one explanation for the difficulties that ABI users have with frequency discrimination (Shannon et al., 1990), and they illustrate the challenges for presenting frequency-specific cues in this neuroprosthesis.

Microfabricated electrode arrays for the ABI

Thin film electrode technology allows for miniaturization and increasing the density of electrode sites (Allitt et al., 2016; Rodger et al., 2008). The thin-film array tested here, using the technology described previously (Guex et al., 2015), accommodates 20 microelectrodes within 2 mm², an electrode density much higher than that of clinical ABIs. Its success was demonstrated by the fact that activity was evoked by nearly all (> 93%) of the electrodes. Given that the rat CN is about 1/16 the size of the human auditory brainstem (Rosahl et al., 2013), these results suggest that a clinical ABI incorporating a much larger number of electrodes could be effective in evoking responses in humans. However, such microelectrodes (< 1 mm²) will need to be optimized against safe charge injection and long-term in vivo use (McCreery et al., 1990).

The surface of neural tissues is usually curvilinear, so we used a strategy to increase the proximity of the electrodes by the use of thin and plastic materials (Baek et al., 2014; Kim et al., 2010; Minev et al., 2015). Furthermore, matching not only the contour and static mechanical response of the tissue but also its dynamics (tissues constantly move due to respiration and cardiac pulsation) opens the path to implants that are better biointegrated in the long term (Minev
et al., 2015). The array used here is conformable, providing close contact of the ABI microelectrodes with the curved surface of the DCN (Fig. 2.1A, Guex et al., 2015). This characteristic may produce the observed low and uniform pattern of thresholds across the DCN's long lateromedial surface (Figs. 2.2E), although this conclusion must be tempered somewhat since: 1) the responses were not explicitly compared to those of a rigid array, and 2) stimulation voltage levels were reported and the exact stimulation current levels are not known. In some human subjects, there is an uneven pattern of response in which the electrodes at the center of the array have lower thresholds than those at more peripheral locations (Barber et al., 2017). This pattern may result from the human ABI array being less conformable so that only the center electrodes are close to the neural tissue. Unfortunately, the higher currents necessary for auditory responses at the periphery of the array may be responsible for activation of other nearby structures in the brainstem. In fact, side effects are a common problem in adult ABI users (Herrmann et al., 2015; McCreery et al., 1998; Ramsden et al., 2016; Wilkinson et al., 2016) and require that such electrodes be turned off during programming. Thus, conformable arrays for a next-generation human ABI may be an important step toward fewer side effects and thus more useable electrodes to provide tonotopy.

**Tonotopy from surface stimulation of the cochlear nucleus**

A challenge for the ABI is to provide tonotopic patterns of response from surface stimulation. Unfortunately, even with a conformable array hosting small electrodes, the unprocessed responses to electric stimulation were not as narrow as those evoked by pure-tone acoustic stimulation (Fig. 2.3A) or by microstimulation within the CN (Mauger et al., 2010; McCreery, 2008; Shivdasani et al., 2008). Processed responses did have larger ranges of CFs
and narrower widths. The first processing used, subtraction of common activity, ignores and perhaps even distorts temporal patterns of response. For that reason the second processing, subtraction of the late phase of response, has its own advantages to reveal the contributions of tonotopically organized neurons such as the DCN principal cells. These cells are almost certainly activated by the array placements used here, because they are located in layer II, just a few hundred micrometers below the surface (Willard et al., 1983). Their tonotopic patterns are likely to be projected centrally into the IC. Future stimulus paradigms might be designed to increase the activity in the pathways providing tonotopic cues.

There are important caveats in extending our animal results to the human ABI. The animal results were obtained in preparations having functional auditory nerve fibers, whereas most ABI users have none. Conceivably, the nerve fibers could have responded to the electric stimulation and carried the response to distant parts of the CN, blurring the tonotopy. In addition, long-term hearing impairment, such as that in NF2, causes plastic changes in the CN (Ryugo, 2015; Salvi et al., 2000), and these effects were not present in our normal-hearing animals. Finally, there may be functional differences between animals and humans, since the DCN is: 1) layered in animals but appears to be unlayered in humans, and 2) contains in animals a prominent type of neuron called granule cells that may be lacking in humans (Baizer et al., 2014).
Figure 2.7: Combined subtraction of common activity and late phase of response shows the largest shifts of CFs.

Plots of CF (A) and width (B) after separate and combined subtractions as a function of lateromedial position or tonotopic axis position (inset plots).

**ABI stimulation of DCN vs. VCN**

In our study, we placed the electrode array on a clearly visible DCN; thus, components of the IC neural response almost certainly result from DCN principal cells. One idea is that other, non-tonotopic components, as well as the eABR, result from stimulation of the VCN, which is
adjacent and shares a common boundary with the DCN (Fig. 2.1B). In support of this idea, the ABR, which is the response to sound stimuli, is generated by VCN bushy cells and their pathways (Melcher et al., 1996). If the eABR is also generated by the VCN, it would explain its distinct spatial variation compared to the IC neural responses found in the present study (Fig. 2.2F). The identity of the generators of the eABR is an important clinical issue, because the eABR is the functional marker used to adjust the position of the ABI electrode array in humans (the surface of the CN is not visible to the surgeon). Bushy cells are thought to be part of a neural pathway mediating binaural localization of sound sources, having excellent phase locking and projections into centers in the superior olivary complex where neurons are sensitive to interaural level and interaural time differences (Smith et al., 1993). However, their contributions to speech perception are unknown. Bushy cells might be responsible for the late-phase IC neural responses (Fig. 2.5B) because their ascending pathway includes a synapse before arriving at the IC (Cant and Benson, 2003). The late-phase responses could also be mediated by VCN “T” stellate cells that project to IC via thin, slow-conducting axons. These possibilities are consistent with microstimulation within the VCN that results in a late phase of activity (Shivdasani et al., 2008), and with lesion studies in mice showing the early phase is eliminated yet the late phase remains after section of the DCN’s output pathway (McInturff et al., 2018). Additional contributors to the late phase could be less tonotopically organized pathways beginning within the DCN (perhaps involving parallel fibers or the auditory nerve fibers that remain in our preparation).
Figure 2.8: Response measures for inter-electrode distances of 400 vs. 800 µm.

A) Plot of CF dependence for pairs with the three orientations (-60, 0 and 60 degrees, as shown in Fig. 2.1C). B) Plot of CF dependence for 400 µm and 800 µm electrode distance (also shown in Fig. 2.1C). C-E) Effect of distance and angle of the stimulation pair on measured best threshold (D), width of activation (E) and normalized spike count (F). Significant differences (p-value < 0.05) are indicated with a star.

Electrode spacing, angle, and monopolar vs. bipolar stimulation

The dense packing of electrodes on the array allowed us to compare closely spaced vs. more distantly spaced electrodes in a bipolar pair. The distantly spaced pairs evoked a less tonotopic pattern of responses and also broadened the width of activation. These results suggest that if bipolar stimulation were used in the clinical ABI, reducing the inter-electrodes distance might improve tonotopic representations. For the rat CN, even the distantly spaced pairs (800
µm) are still closer than for the spacing in the available clinical ABI electrode array (about 1 mm), suggesting that the tonotopy provided by the clinical device is particularly low.

The DCN principal neurons have axons and dendrites travelling perpendicular to the ventrolateral-to-dorsomedial tonotopic axis, and thus may be activated differently with current pulses applied with various orientations. As shown in the data obtained with the larger inter-electrode distance (800 µm), the orientation of the pairs may influence the shaping of the current field and volume of neural tissue that is activated in CN (Bonham et al., 2008), resulting in varying tonotopic organization of responses. Electrode pair orientations with negative angles (close to orthogonal to the tonotopic axis after combined subtraction) had two differences from other orientations: they resulted in activation of a larger range of recording sites along the IC’s tonotopic axis, and the widths of the activated areas were narrower. Perhaps this results from the long axis of the electric field being perpendicular to the DCN’s tonotopic axis. If bipolar stimulation were used in the clinical ABI, pair orientation should be chosen to be perpendicular to the tonotopic axis.

Our data indicate that bipolar electrode configurations alone do not usually result in narrower receptive fields compared to the monopolar condition (Figs. 2.9, 2.10D). Such a result might suggest that the volume of tissue activated by the electric field is roughly similar for the two conditions. One possibility is that for bipolar configuration, much of the current path is directly from one electrode to the other through the superficially-most tissue of layer I, and thus avoiding the principal cells. This idea is consistent with our finding that spike rates in response to monopolar stimulation were on average higher than to bipolar stimulation. Another possibility is that in biphasic stimulation both electrodes of a bipolar pair act as cathodes/anodes. Contrary to studies comparing monopolar and bipolar stimulation in the cochlea (Allitt et al., 2016;
Bonham et al., 2008; Middlebrooks et al., 2007; Snyder et al., 2008; Zhu et al., 2012), the threshold obtained with bipolar stimulation is not higher than with monopolar stimulation. In the human ABI, bipolar stimulation is used intra-operatively to assist the positioning of the ABI array, but monopolar stimulation is used to minimize power consumption (Herrmann et al., 2015).

Acknowledgements

We thank Evan Foss for technical assistance. Preliminary versions of this work were presented at the Cell Symposium Engineering the Brain, Chicago, 2015, at the Materials Research Society spring meeting, 2017, and are contained in the EPFL Ph.D. thesis of Amélie Guex.

Funding: Supported by the Bertarelli Foundation, and by NIDCD grant DC01089.
Figure 2.9: Similarities of responses to monopolar and bipolar stimulation in an individual rat (Rat #11).

A) IC response maps (raw data) to bipolar electrical stimulation with the left column of plots showing five different bipolar pairs selected along the lateromedial axis and the middle and right columns of plots showing the response to monopolar stimulating using the corresponding individual electrodes. Red arrows and numerals indicate centroids of activity. B) CF matrices (raw data) for bipolar stimulation (top) or monopolar stimulation (bottom). C) CF matrices (after combined subtraction) for bipolar stimulation (top) or monopolar stimulation (bottom). Thin dashed lines represent electrode pairs producing insufficient response to estimate the CF.
Figure 2.10: Similarities of responses to monopolar and bipolar stimulation in all rats.

CF dependence of (A) raw responses and (B) after combined subtraction, for monopolar and bipolar stimulation. Bar graphs of best threshold (C), width of activation (D) and normalized spike rate (E) for monopolar and bipolar stimulations, before (raw data) and after combined subtraction.
CHAPTER 3.

SUPERIOR TEMPORAL RESOLUTION OF CHRONOS VERSUS CHANNELRHODOPSIN-2 IN AN OPTOGENETIC MODEL OF THE AUDITORY BRAINSTEM IMPLANT


* denotes equal contribution

Abstract

Contemporary auditory brainstem implant (ABI) performance is limited by reliance on electrical neurostimulation with its accompanying channel cross talk and current spread to non-auditory neurons. A new generation ABI based on optogenetic-technology may ameliorate limitations fundamental to electrical stimulation. The most widely studied opsin is channelrhodopsin-2 (ChR2); however, its relatively slow kinetic properties may prevent the encoding of auditory information at high stimulation rates. In the present study, we compare to ChR2 the temporal resolution of light-evoked responses of a recently developed fast opsin, Chronos in a murine ABI model. Viral mediated gene transfer via a posterolateral craniotomy was used to express Chronos or ChR2 in the mouse nucleus (CN). Following a four to eight week incubation period, blue light (473 nm) was delivered via an optical fiber placed directly on the surface of the infected CN, and neural activity was recorded in the contralateral inferior colliculus (IC). Both ChR2 and Chronos evoked sustained responses to all stimuli, even at high pulse rates. In addition, optical stimulation evoked excitatory responses throughout the tonotopic
axis of the IC. Synchrony of the light-evoked response to stimulus rates of 14-448 pulses/s was higher in Chronos compared to ChR2 mice (p<0.05 at 56, 168, and 224 pulses/s). Our results demonstrate that Chronos has the ability to drive the auditory system at higher stimulation rates than ChR2 and may be a more ideal opsin for manipulation of auditory pathways in future optogenetic-based neuroprostheses.

Introduction

The cochlear implant (CI) is the most successful of neuroprostheses, and provides meaningful auditory benefits to pediatric and adult patients with severe to profound hearing loss. In the past 50 years, over 300,000 individuals worldwide have received the CI (NIDCD, 2014). Over this period, CI technology has evolved from a crude single channel implant to a multi-channel auditory neurostimulator providing sound and speech perception to the majority of deaf users. Cochlear implants have proven highly beneficial for several etiologies of hearing loss, including genetic causes of deafness (Vivero et al., 2010). The recent Lasker Award highlights the development of the CI and illustrates the profound success of this device and its positive impact on society (Williams, 2013). However, there is a small subset of deaf individuals who will not benefit from the CI due to 1) a small or absent cochlea, 2) a small or absent auditory nerve, or 3) injury or scarring of the inner ear or auditory nerve secondary to meningitis, trauma, or tumor, such as bilateral vestibular schwannomas that arise from Neurofibromatosis-2 (NF-2) (Asthagiri et al., 2009). An auditory brainstem implant (ABI) is an option to provide hearing sensations in these patients who are not candidates for the CI due to these considerations. More than 1,000 patients worldwide have been implanted with an ABI (Lin et al., 2012). The ABI bypasses the damaged or absent cochlea and auditory nerve to transmit electrical stimuli to the cochlear nucleus (CN) in the brainstem (Hitselberger et al., 1984; Sennaroglu et al., 2009).
Hearing outcomes of ABI users are highly variable across similar cohorts of patients (Colletti et al., 2012; Colletti and Shannon, 2005; Nevison et al., 2002), and overall performance of ABI users lags behind that seen for CI users. Further, many ABI users experience side effects, such as facial pain, tingling, and twitching, as well as dizziness, due to activation of non-auditory neurons (Colletti et al., 2010). One possible explanation for limited outcomes and side effects may be the spread of electric current (Eisen and Franck, 2005; Nardo et al., 2008; Venter and Hanekom, 2014). One approach to improve speech perception is to increase the number of electrode channels. However, due to current spread, this may result in channel cross talk (Boëx et al., 2003; Karg et al., 2013; Qazi et al., 2013).

Optical stimulation of the nervous system is now being used as a novel stimulus paradigm in research laboratories. For the central auditory system, light-based activation offers a theoretical advantage over traditional electric-based neural stimulation as focused light may be able to excite a select set of neurons, increasing the density of independent stimulation channels while reducing the unintended consequence of current spread (Fu and Nogaki, 2005; Fu et al., 1998). These properties could address the limitations seen with the electrically based ABI. Over the past decade, infrared neural stimulation (INS) of the auditory system has been investigated as an alternative means to stimulate neurons; however, INS may have limited applications in the central auditory system. Recent efforts employing INS in the central auditory system have failed to elicit an evoked response in a deafened animal model, limiting its potential clinical utility (Verma et al., 2014).

In contrast to INS, optogenetics uses light from the visible spectrum to stimulate the nervous system, and it has been used to investigate a host of neural systems (Ayling et al., 2009; Boyden et al., 2005; Huff et al., 2013; Rolls et al., 2011). Viral-mediated infection is a common
approach to deliver genes encoding for microbial opsins, light-gated transmembrane channels that enable neurons to respond to optical stimulation. ChR2 is the most widely used opsin in neuroscience, (Bernstein et al., 2008; Boyden et al., 2005; Chow et al., 2010; Han and Boyden, 2007; Zhang et al., 2006), however, only a few recent studies have applied optogenetics to the auditory system. (Hernandez et al., 2014; Shimano et al., 2013) Shimano et al. introduced ChR2 into the CN and demonstrated light-evoked increases in auditory neural activity locally in the CN. Building on the results of Shimano et al., we previously showed optogenetic stimulation of the CN results in activation of the upstream auditory pathway, including the inferior colliculus and auditory cortex (Darrow et al., 2013). In a recent optogenetics study of the peripheral auditory system, transgenic mice expressing ChR2 in spiral ganglion neurons (SGN) of the cochlea showed neural responses in SGNs themselves and in upstream neurons in the central nucleus of the Inferior Colliculus (ICc) (Hernandez et al., 2014). Overall, these studies demonstrate that optogenetics can be used to activate the auditory system from the periphery throughout the central pathway.

One unique property of the auditory system is its capability of providing a highly synchronous response with the rapidly varying features of an acoustic waveform, a property necessary to encode the also rapidly varying characteristics of speech. Original studies of the kinetics of ChR2 and its variants (Boyden et al., 2005; Zhang et al., 2006) suggest that ChR2 may be too slow for optimal function in the auditory system (Darrow et al., 2013). Over the last several years, a host of new opsins have become available with variable activation thresholds, wavelengths of stimulation, and, most importantly, kinetic properties (Yizhar et al., 2011). One of the most recently developed opsins, Chronos (Klapoetke et al., 2014), appears to have faster kinetic properties that may better suited to temporal characteristics of the auditory system.
Herein, we compare the temporal characteristics of ChR2 and Chronos in a translational murine ABI model.

Figure 3.1: Surgical approach to the DCN for opsin injection and IC for neural recording with a multichannel electrode probe.

A: The skin and muscle is retracted laterally to expose the lambda and the coronal suture lines. B: Placement of NeuroNexus recording probe into the IC in a vertical direction. C: Left-sided posterior craniotomy and partial cerebellar aspiration have been performed and a 400 μm diameter optical fiber mounted on a micromanipulator is introduced through the craniotomy and onto the CN surface. D: Schematic representation of the recording probe positioned along the tonotopic axis of the central nucleus of the inferior colliculus (ICc) so that each of the 16 recording electrodes (spaced 50 μm apart) records a different characteristic frequency.
Methods

Animal Protocol

All experimental procedures were performed in accordance with the National Institutes of Health guidelines for the care and use of laboratory animals as well as the approved animal care and use protocols at the Massachusetts Eye & Ear Infirmary, Boston, MA.

Surgical Exposure of the Dorsal Cochlear Nucleus

Methods to expose the dorsal cochlear nucleus (DCN) have been described (Kozin et al., 2014). Direct exposure of the DCN, rather than stereotaxic injection, was used to minimize the chances of missing the desired injection site in addition to replicating the potential surgical approach used during human ABI operations. Normal hearing CBA/CaJ mice aged 4–6 weeks were anesthetized with xylazine (10 mg/kg) and ketamine (100 mg/kg) via an intraperitoneal (IP) administration. Following anesthesia, the overlying the scalp was exposed to provide unobstructed access to the surgical site. The mouse was placed in a Kopf small animal stereotaxic holder (Tujunga, CA), and held in place by a snout clamp. The left parietal, interparietal, and occipital bones of the skull are exposed and rongeurs are used to make a craniotomy over the interparietal bone, left of midline, ~2 mm caudal to the lambda suture line. Following craniotomy, using a 5 French suction, aspiration of lateral-most portion of the left cerebellum reveals the underlying DCN (Fig. 3.1).

Pressure Microinjection for Gene Transfer

After the DCN was clearly visible, pressure microinjections are made into the DCN using a 5 µl Hamilton syringe. Between 1.5 and 2.0 µl of adeno-associated virus with ChR2 (AAV2.8-
ChR2 fused with GFP or mCherry and CAG promoter, courtesy of Dr. Edward Boyden’s lab) or Chronos were infused over a 2 to 4 minute period. (Plasmid for Chronos was provided by Dr. Edward Boyden, Massachusetts Institute of Technology, Cambridge, MA. Amplification took place at the Boston Children’s Viral Core, Boston, MA. Measured titers of Chronos were $1.21 \times 10^{14}$ GC/ml). Immediately following injection, the incision was closed and the scalp was sutured. Four additional mice were used as either ‘sham’ or control cases. These included sham-injected mice ($n=2$) that underwent the exact surgical protocol as AAV injected mice, including injection of saline into the brainstem over 2-4 minutes, and control mice ($n=2$), which had no history of manipulation.

Re-exposure of Dorsal Cochlear Nucleus and Exposure of Contralateral Inferior Colliculus

After a four to eight week survival time to allow expression of either ChR2 or Chronos, the mice were prepared for acute surgery to characterize responses to optical stimulation. Mice were re-anesthetized and underwent the above-described surgical procedure for performing the craniotomy and the cerebellar aspiration for direct visualization of the dorsal cochlear nucleus. After the injected region was re-exposed for optical stimulation, a craniotomy was made over the right IC and the exposed brain surface was covered with high-viscosity silicon oil. During the course of physiological recordings, the core body temperature of the animal was maintained at 36.8°C with a homeothermic blanket system.

Optical Stimulation

Blue light (473 nm) stimuli were produced by a laser (BL473T-100FC, Shanghai Laser & Optics Century Co.) targeting the peak wavelength sensitivities of ChR2 (~475 nm) and Chronos
(~500 nm) (Klapoetke et al. 2014). Light stimuli were delivered via an optical fiber (400 µm diameter) that was held in place by a micromanipulator and placed directly on the exposed surface of the DCN (Fig. 3.1). Pulses of 1 ms duration were presented at pseudorandomized rates from 14-448 pulses/s for train durations of 300 or 500 ms followed by 300 or 500 ms of no stimulation, respectively. For both protocols, either 50 or 80 trials were presented at light intensity ranging from 0 to 13 mW. The laser was calibrated by positioning the optical fiber ~2 mm from a standard Si photodiode power sensor (9.5 mm diameter, Thorlabs S121C) connected to a USB power and energy meter interface (Thorlabs PM100USB). The voltage command parameters were systematically varied based on the range of pulse amplitudes, widths, and rates used experimentally. Resulting measured laser intensity (radiant exposure, mW) was calibrated to the voltage input.

**Contralateral Inferior Colliculus Recordings**

Multiunit recordings were made from the central nucleus of the IC using a penetrating 16-channel linear silicone probe (NeuroNexus Technologies, Fig. 3.1). The position of the recording probe was first inserted perpendicular to the exposed surface above the ICc, approximately ~1 mm lateral to the midline and immediately caudal to the transverse sinus. The probe was repositioned until a complete tonotopic map across the recording channels was obtained (Guo et al., 2012; Malmierca et al., 1993) using acoustic frequencies from 8 to 45.25 kHz in 0.5 octave steps and from 0-80 dB in 10 dB steps, using 20 ms duration tone bursts with a repetition rate of 10 bursts/s. Injected mice had acoustic thresholds (<40 dB SPL) comparable to non-injected mice. Raw voltage signals were band-pass filtered (0.3-3 kHz, 5 pole) and
sampled at 25 kHz. Common mode rejection was performed across all 16 channels and then the signal was digitally filtered (zero-phase butterworth band-pass filter, 0.5-3 kHz, 5-pole).

To compute average firing rate, spike count was computed over the pulse train duration. Spontaneous firing rate was computed from spikes collected during the no stimulation period. Driven rate was computed by subtracting the spontaneous rate from the firing rate during the stimulus. Paired t-tests were performed between the firing rates and spontaneous rates from every trial for each electrode and stimulus intensity. For each electrode-intensity combination in which \( p > 0.01 \), the driven rate was assigned to be 0 spikes/second. Driven rate-intensity curves were generated from the average of driven rate across all electrodes collected during each stimulus intensity presented at 28 pulses/s and normalized to the highest driven rate evoked by any stimulus intensity. To investigate temporal properties of laser-evoked spiking we computed the synchronization index (SI, magnitude of the vector of averaged spikes collected during the period between stimulus pulses; the SI varies between 0 (no synchronization) and 1 (all spikes occurring exactly at the same phase of the stimulus period) (Dynes and Delgutte, 1992). For all stimulus electrode-intensities in which driven rate was calculated to be zero spikes/second, SI was also set to zero.

**Histology and Immunohistochemistry**

Following conclusion of experiments, mice were euthanized with an overdose of ketamine and perfused with normal saline followed by 4% paraformaldehyde. Brainstems were extracted from the skull and post-fixed for 2 hours. Brainstems were cryoprotected in 30% sucrose for 48 hours, and then sectioned using a cryostat into 30-60 µm sections. Before the staining procedure, sections were allowed to dry at room temperature and then rehydrated in PBS.
for 10 min. After washing with PBS, tissue was permeabilized and blocked with blocking solution (0.3% Triton X-100, 15% heat inactivated goat or donkey serum in PBS) for 1 hour. Visualization of cell nuclei was enabled with 4,6-diamidino-2-phenylindole (DAPI), (Vector Laboratories). Staining was analyzed with epifluorescence microscopy (Axioskop2 Mot Axiocam, Zeiss) and confocal microscopy (Leica).

Results

Expression of Chronos and ChR2 in the Cochlear Nucleus

Opsin-linked immunofluorescence demonstrated Chronos or ChR2 transfer throughout the dorsal cochlear nucleus (DCN) and in the ventral cochlear nucleus (VCN). Chronos-GFP immunofluorescence appeared in an array of cell types, including morphologies consistent with DCN fusiform cells (red arrow, Fig. 3.2B), giant cells, and cartwheel cells. We did not directly investigate anterograde labeling of axons in Chronos cases, but such labeling has been observed previously in ChR2 cases, which employed an identical AAV serotype and promoter (Darrow et al., 2014).

Figure 3.2: Chronos expression in the cochlear nucleus.

A: Mosaic confocal image showing Chronos-GFP expression within the dorsal cochlear nucleus (DCN) and ventral cochlear nucleus (VCN). B: Confocal image of the DCN demonstrates Chronos-GFP expression within a fusiform cell (red arrow) and in other neuronal populations. DAPI demonstrates cell nuclei.
Neural activation in the IC was evoked by light stimulation of the CN. Dot raster and peristimulus time (PST) histogram plots from one electrode of a Chronos mouse are shown in Figure 3.3. High, sustained rates of firing are observed during the light pulse trains and low rates of spontaneous firing are observed when there are no stimuli (second half of traces). For the pulse train of 448 pulses/s (Fig. 3.3B), the PST histogram shows that the high initial firing adapted to steady-state firing over the course of the 500-ms pulse train. Even during the adapted portion of the evoked response, the driven rates were sustained and above spontaneous firing. This adaptation was not present for the low rate (Fig. 3.3A), even though the overall firing rate was greater.

The temporal pattern of evoked neural activity in the IC depends on the stimulus pulse rate. For the train of pulses at 28 pulses/s (Fig. 3.3A), evoked activity is synchronized to the pulses, whereas for the train of pulses at 448 pulses/s (Fig. 3.3B), evoked activity, though high, appears to be less synchronized. The average SI values are plotted against stimulus rate for Chronos and ChR2 cases in Figure 3.4A. For both opsins, there is a decline in SI with increasing pulse rate. For all rates, recordings from Chronos mice had higher SI than those from ChR2 mice. These differences were significant at rates of 56, 168 and 224 pulses/s (Fig. 3.4A, asterisks).

The decline in synchrony with increasing pulse rates was not due to a decline in firing rate (Figs. 3.4, 3.5) and driven rate was significant at all tested stimulus rates (14-448 pulses/s, Fig. 3.4B). Even at high pulse rates (e.g. 448 pulses/s), evoked activity was significantly above spontaneous (paired t-test), high (average 76 spikes/s), and sustained over the entire stimulus...
pulse train (Fig. 3.3B) despite the compromised kinetics of both Chronos and ChR2. Further, the firing rates of mice injected with Chronos or ChR-2 were increasing, monotonic functions of laser intensity, with evidence of incomplete saturation at the highest tested intensities (Fig. 3.5A). At these intensities, average firing rates were significantly lower (p < 0.001) for ChR2 versus Chronos, average rate 55 spikes/s for ChR2 vs. 127 spikes/s for Chronos (~12 mW, stimulus rate = 28 pulses/s).

Figure 3.3: ICc neural activity evoked by light pulses (Chronos injected mouse).

A: Responses at a low stimulus rate (28 pulses/s) elicit synchronized spikes (dots on raster plot at top). Bottom plot shows PST histogram for the same data. B: Responses at a high stimulus rate (448 pulses/s) show less synchrony (synchronization index indicated on the plots). In addition, the PST histogram shows the prominent rate adaptation during the pulse train. Driven rates (an average over the 500 ms duration pulse train) are indicated on the plots.
Figure 3.4: Synchronization index and driven rate as a function of pulse rate.
A: Average SI (at ~12 mW) is significantly greater for Chronos than ChR2 at stimulus rates of 56, 168 and 224 pulses/s (* = p < 0.05, two-sample t-test). B: Average maximal driven rates over all 16 electrodes.

Spatial Pattern of Response for Optogenetic-based Stimulation of the Cochlear Nucleus

There was variability in the spatial pattern of responses from case to case. Figure 3.5B shows a case where nearly all electrodes recorded laser-evoked activity for stimulus levels at and above threshold (driven rate > 0 spikes/s, indicated by non-dark blue coloring). Highest spike rates were observed on electrodes 2–8. Other cases (data not shown) had variable patterns of neural activation: of the 16 electrodes, the average number activated across all laser intensities at and above threshold (stimulus rate = 28 pulses/second) was 12.4 with a range of 9-14.9 electrodes for Chronos (n=4 mice) and 9.2 with a range of 4.6-14 electrodes for ChR2 (n=4 mice). Our data set is insufficient to correlate the opsin staining pattern with the spatial pattern of responses. Sham and control mice showed no response to optical stimulation (one example shown in Fig. 3.5C).

Discussion
**Chronos versus ChR2 for Light-Evoked Activation of the Auditory System**

Our study is the first to characterize the temporal properties of opsins in the central auditory system. Of all presently studied opsins, Chronos has the fastest on/off kinetics based on firing rates in *in vitro* studies (Klapoetke *et al.*, 2014). For *in vivo* stimulation of the CN, we found that Chronos enabled better neural synchrony compared to ChR2. For both opsins, there was a decline in SI with increasing pulse rates; however, the decline was more pronounced for ChR2. Significant differences in the SI of evoked activity between the two opsins were found at 56, 168 and 224 pulses/s. Synchronous activity at these pulse rates are clinically significant because contemporary clinical ABI processors employing the SPEAK sound processing strategy use pulse rates of 250 pulses/s; this pulse rate is a carrier that works well for neural encoding of amplitude modulation (McCreery *et al.*, 2013).

Although the temporal properties of optogenetic responses have not been characterized previously, the responses of IC neurons to electrical stimulation of the cochlea have previously been reported. These responses also decline with increasing pulse rates, and their synchrony measures are comparable to the optogenetic responses reported here (Middlebrooks and Snyder, 2010; Snyder *et al.*, 1995). In addition, synchrony is even higher in IC recordings from awake animals under comparable stimulus conditions (Chung *et al.*, 2014). In those studies, differences among types of units were observed, with some units able to fire synchronously to high rates (> 100 pulses/s) whereas others unable to synchronize to these rates. In the present study, single recording sites probably sampled from a combination of these unit types, future studies could employ spike sorting to examine the synchronization limit of single units to optical stimulation.

During optogenetic stimulation, even for pulse rates for which responses were non-synchronized, driven rates were substantial (Fig. 3.3B). Such driven rates will signal the
presence of a stimulus even though there is little synchrony to the fine time structure of the stimulus. Further, since rate-intensity curves (Fig. 3.5B) suggest that the entire dynamic range of the response was not captured due to the limitations of our laser, a higher-intensity stimulus would likely produce even higher driven rates. However, the level at which potential light-induced damage manifests is not clear.

**Limitations of Viral-mediated Gene Transfer**

There are inherent limitations to viral mediated gene transfer in the central nervous system that may have influenced our results. We previously demonstrated that expression of opsins as a result of viral mediated gene transfer is variable from case to case (Darrow et al., 2014; Darrow et al., 2013). Consequently, the locations and numbers of activated cells differed somewhat between Chronos and ChR2 cases. This may affect synchrony because, at least for responses to acoustic stimuli, different types of neurons have different temporal characteristics (Young, 1984). We believe that at least some of the excitatory responses observed here may have been mediated by fusiform cells of the DCN because in almost all cases they express Chronos or ChR2 and because they project directly to the IC (Darrow et al., 2014; Oliver, 1984; Oliver and Morest, 1984). Perhaps the slow kinetics of ChR2 in these cells accounts for both the low SI and lower driven rates observed in our study. Regardless of the exact etiology, the performance of Chronos is superior in both driven rate and SI.
Figure 3.5: Neural activity as a function of stimulus intensity and recording position.

A: Plot of normalized driven rate across all electrodes as a function of laser intensity for the two opsins. (Stimulus rate = 28 pulses/s) B: Response map showing driven rate as a function of electrode (position in IC, see Fig. 3.1D) number and laser intensity. (Chronos injected mouse, stimulus rate = 56 pulses/s) C: Control mouse response map showing no response evoked by light. (Stimulus rate = 28 pulses/s.)

Translational Models and Future Optogenetics Research in the Central Auditory System

In addition to examining the temporal properties of Chronos and ChR2, our study demonstrates a feasible translational approach for gene transfer of opsins to the CN. Specifically, our surgical approach in the murine model allows for visualization of the DCN and inoculation with a viral gene transcript. In many respects, our model is analogous to surgical approach of the human ABI placement. While viral mediated gene transfer has inherent weaknesses and risks, numerous FDA-approved gene therapy studies are ongoing and may be a potential tool for gene
transfer to the CN of humans.

Further, our study represents the beginning of a path toward the particular opsin chosen for an eventual prosthesis. We define the ideal characteristics of an “auditory opsin” for use in a neuroprosthesis: 1) fast temporal kinetics to encode speech information, 2) low activation threshold to decrease energy requirements from illumination powered by an external source, and 3) a promoter that enables tissue-specificity using virally-mediated gene transduction to CN (to minimize non-specific expression of opsins). Looking forward, our study raises several questions: What is the ideal gene transfer approach for delivery of opsins to the auditory system? What is the long-term safety profile of opsins, and will there be any deleterious effects from their presence? Finally, can optogenetic-based stimulation function as replacement to electrical stimulation, and can it be used as an adjunct? The answers to these questions, both in the central and peripheral auditory systems, remain to be seen and should be the focus of future studies.

Conclusion

Previous studies have demonstrated the feasibility of optogenetic stimulation for light-based activation of the central auditory system. Currently, the most widely used opsin in neuroscience is ChR2, however, it may not possess the temporal properties necessary to encode auditory information. We find in an ABI animal model that Chronos has significantly improved kinetic properties compared to ChR2. These studies highlight the need to further examine and identify the ideal opsin that can support the high stimulation rates needed for the transfer of temporal cues along the auditory pathways. Future studies may seek to design opsins optimized for a new generation ABI based on light.
Acknowledgements

Preliminary results of this study were presented at the Association for Research in Otolaryngology Midwinter Meeting, February 2013 and 2014. This work was supported by a Fondation Bertarelli grant (DJL, MCB, and SL), a MED-EL grant (DJL), and National Institutes of Health Grants DC01089 (MCB), T32 DC000038 (AEH), T32 DC000020 (EDK).
HEARING THE LIGHT: NEURAL AND PERCEPTUAL ENCODING OF OPTOGENETIC STIMULATION IN THE CENTRAL AUDITORY PATHWAY

Wei Guo, Ariel E. Hight, Jenny X. Chen, Nathan C. Klapoetke, Kenneth E. Hancock, Barbara G. Shinn-Cunningham, Edward S. Boyden, Daniel J. Lee, & Daniel B. Polley

ABSTRACT

Optogenetics provides a means to dissect the organization and function of neural circuits. Optogenetics also offers the translational promise of restoring sensation, enabling movement or supplanting abnormal activity patterns in pathological brain circuits. However, the inherent sluggishness of evoked photocurrents in conventional channelrhodopsins has hampered the development of optoprostheses that adequately mimic the rate and timing of natural spike patterning. Here, we explore the feasibility and limitations of a central auditory optoprosthesys by photoactivating mouse auditory midbrain neurons that either express channelrhodopsin-2 (ChR2) or Chronos, a channelrhodopsin with ultra-fast channel kinetics. Chronos-mediated spike fidelity surpassed ChR2 and natural acoustic stimulation to support a superior code for the detection and discrimination of rapid pulse trains. Interestingly, this midbrain coding advantage did not translate to a perceptual advantage, as behavioral detection of midbrain activation was equivalent with both opsins. Auditory cortex recordings revealed that the precisely synchronized midbrain responses had been converted to a simplified rate code that was indistinguishable between opsins and less robust overall than acoustic stimulation. These findings demonstrate the temporal
coding benefits that can be realized with next-generation channelrhodopsins, but also highlight the challenge of inducing variegated patterns of forebrain spiking activity that support adaptive perception and behavior.

INTRODUCTION

For individuals with profound peripheral nerve degeneration, the only available treatment avenue for sensory restoration lies in delivering patterned electrical stimulation at early stages of central sensory processing. Although central prostheses have been used in human patients for over 50 years, they generally provide only a rudimentary sensory awareness. For example, translating spatial forms in the visual field to retinotopically organized microstimulation of primary visual cortex provides an awareness of object location, but does not support form discrimination (Schmidt et al. 1996; Dobelle et al. 2002). Similarly, hundreds of deaf individuals have gained an awareness of environmental sound through auditory brainstem or midbrain implants. However, discrimination of complex signals such as speech is generally quite poor (Otto et al. 2002; Nevison et al. 2002; Otto et al. 2008; Lim et al. 2009), with a few notable exceptions (Colletti et al. 2009; Shannon et al. 2012).

Interestingly, prostheses that deliver patterned electrical stimulation to retinal ganglion or spiral ganglion processes generally provide superior outcomes than stimulation of low-level brain areas (for review see Shannon 2012, Weiland et al. 2014). The abrupt performance drop associated with prosthetic devices that stimulate the brain rather than peripheral nerves may arise from more demanding surgical placement and electrical signal processing; yet the foremost cause arguably lies in the enormous increase in the complexity of coding and processing within brain networks themselves. Brain circuit organization varies widely across regions but most share a
common logic that includes interconnected afferent recipient neurons, interneurons, feedback neurons, projection neurons and a host of glial cells that modulate chemical and electrical neurotransmission. Whereas electrical microstimulation indiscriminately activates multiple elements of these circuits, the use of genetically encoded light-activated ion channels (i.e., optogenetics) could provide the means to pinpoint stimulation to specific nodes within these circuits (for review see Bernstein et al. 2011, Fenno et al. 2011).

A host of ethical, engineering, and biological hurdles stand between the current implementation of optogenetic technologies in basic science research and the targeted delivery of channelrhodopsins to specific cell types in human patients. One fundamental problem lies in the fact that light-activated photocurrents are sluggish due to inherently slow channel kinetics in channelrhodopsins, which reduces their ability to deliver long lasting, high-frequency neural stimulation (e.g., above 40Hz, Boyden et al. 2005, Lin et al. 2009). This limitation is particularly problematic for any attempt to reconstitute sound representations in early stages of the auditory pathway, where temporal modulations in acoustic signals are normally encoded with submillisecond precision at rates as high as several hundred hertz (Joris et al. 2004; Golding et al. 2012; Carr 2004). Because precise, sustained spike timing is essential for auditory feature encoding, faithful reconstruction of sound representations by optogenetic stimulation also requires an ability to induce precise, fast, and non-adapting patterns of spiking activity.
Recently, a channelrhodopsin nicknamed ‘Chronos’ was identified in the algal species Stigeoclonium helveticum. Using a combination of patch clamp recordings from cultured neurons and acute brain slices, Chronos was revealed to have the fastest channel kinetics described in any naturally occurring or genetically engineered channelrhodopsin (Klapoetke et
al. 2014). The advent of Chronos inspired us to explore the feasibility of an optoprosthesis for the central auditory pathway. The present experiments explore the suitability of inducing actionable auditory percepts by activating early stations of the central auditory pathway with optogenetic stimulation. Our first aim was to determine whether the superior temporal fidelity of Chronos over conventional ChR2 could also be demonstrated in vivo, using the central nucleus of the inferior colliculus (ICc), a central midbrain hub for acoustic signal analysis, as a test bed. We then determined how midbrain temporal coding differences translated to perceptual salience by measuring behavioral detection of acoustic or laser pulses presented at various rates. As a final step, we addressed discrepancies between neural coding in midbrain and behavioral performance by documenting how high-fidelity temporal codes for optogenetic stimulation in the midbrain were transformed in ostensibly disadvantageous ways at the level of the auditory cortex. Collectively, these findings support the feasibility of single channel optoprosthetic implants for basic auditory awareness but underscore the need to develop more sophisticated approaches for multi-channel cell type-specific activation for encoding spectrotemporally complex signals such as speech.
Figure 4.2: Chronos supports a superior neural code for stimulation rate.

(a-d), A PSTH-based approach for single trial pulse rate classification from individual recording sites. (a), Representational templates for each pulse rate are established from 10 randomly selected trials (bottom row) and the remaining 10 trial are then individually assigned to the PSTH template that provides the closest match. Confusion matrices from representative single recording sites illustrate that classification is fairly accurate at low pulse rates but trails off for higher rates with narrowband noise (b) and ChR2 (c) but remains high across the full range of rates tested with Chronos (d). (e), Mean probability of veridical classification (i.e., the upward central diagonal from the confusion matrices) for each stimulation type. Error bars = sem.
Methods

All procedures were approved by the Massachusetts Eye and Ear Infirmary Animal Care and Use Committee and followed the guidelines established by the National Institute of Health for the care and use of laboratory animals. A total of 20 male CBA/CaJ mice were used in this study (10 for in vivo ICc recordings; 10 for behavioral assessments and in vivo Actx recordings).

Virus injection

Adult CBA/CaJ mice aged 8–10 weeks were sedated with isoflurane (5% in oxygen), then anesthetized with ketamine (100mg/kg) and xylazine (10mg/kg). A surgical plane of anesthesia was maintained with supplements of ketamine (50mg/kg) as needed. Throughout the procedure, the animal’s body temperature was kept at around 36.5 °C with a homeothermic blanket system. After numbing the scalp with lidocaine (0.5%), an incision was made along the midline, exposing the skull around the lambdoid suture. A small craniotomy (0.2×0.2mm, with the medial-rostral corner positioned at 0.4mm lateral and 0.1mm caudal to lambda) was made with a scalpel to expose the right inferior colliculus. The dura mater was left intact. Electrophysiological recordings were made to identify the location of the central nucleus (ICc) before virus injection (See acute electrophysiology in the IC). Glass capillary pipettes were pulled and back filled with mineral oil before loading with virus. A motorized stereotaxic injector (Stoelting Co.) was used to inject 0.3–0.5μl of either AAV-CAG-ChR2-mCherry or AAV-Synapsin-Chronos-GFP into the right ICc of the mouse approximately 700μm below the brain surface with an injection rate of 0.05μl/min. The pipette was left in place for an additional 10minutes before withdrawal. The craniotomy was covered with high viscosity silicon oil, and the scalp was sutured back. Mice were allowed to recover for at least 48hours before behavioral
training with NBN and at least 3 weeks before detection was measured with optogenetic stimulation.

**Acute electrophysiology in the IC**

The surgical procedure was similar as described in the previous section. Mice were anesthetized and a craniotomy performed over the right IC. Single-shank multi-channel silicon optrodes (NeuroNexus Technologies) were used to deliver laser pulses and record neural activity (sampled at 24 kHz, digitized at 32 bit, and then band-pass filtered between 300 to 5000Hz with second-order Butterworth filters). Multiunit spike events on each channel were time stamped at threshold crossing (4.5 s.d. above a 10 s running average of the baseline activity, SpikePac, Tucker-Davis Technologies). All recordings were performed in a double-walled sound-attenuating chamber. The ICc was identified according to the dorsal-ventral low-high tonotopic organization as defined by a pseudorandom series of pure tone pips (4–64kHz in 0.1 octave steps, 0–60dB SPL in 5dB steps, 50ms duration with 5ms cosine ramps at the onsets and offsets, 500ms inter-trial intervals) presented to the contralateral ear with a custom-built, calibrated in-ear acoustic system.

Laser pulses (473nm, 1ms pulse width, 1 s total duration, LaserGlow Co.) were presented at various rates (20 to 300Hz, 20Hz steps) to the IC via the optic fiber on the optrode, which was positioned 0.2mm above the topmost recording site. To avoid potential contamination through photoelectric artifacts, threshold crossings during the laser pulse were disregarded. Laser powers were selected to generate suprathreshold responses in the infected tissue on a case-by-case basis, and were generally in the range of 5 to 7mW. For the sake of comparison narrowband noise bursts (filtered from broadband noise stimuli using fourth-order Butterworth filters, 20 kHz center frequency and 0.25 octave bandwidth, 1ms duration, 60dB SPL) were presented at the
same rates via the contralateral in-ear acoustic system. Laser and acoustic stimulation were presented in a pseudorandom order and repeated 20 times each.

**Optic fiber implantation**

Following 2–6 weeks of behavioral testing with acoustic stimuli, mice were anesthetized with ketamine and xylazine, as described previously. An implantable 4mm optic fiber assembly (NeuroNexus NNC fiber) was advanced 0.35mm into the ICc along the previous injection site. The implant was then securely cemented on the skull (C&B Metabond). Mice were allowed to recover for at least 48 hours prior to the continuation of behavioral testing.

**Acute electrophysiology in the auditory cortex**

Following previously described procedures (Guo et al. 2012), ChR2+ and Chronos+ mice were anesthetized with ketamine and xylazine, and a craniotomy was made over the right auditory cortex. The exposed dura was covered with high viscosity silicon oil. Extracellular recordings of multiunit activity were made with tungsten electrodes (FHC Co.) positioned in the middle cortical layers. Acoustic stimuli were delivered to the contralateral ear via a calibrated in-ear acoustic system. Laser stimuli were delivered through the implanted optic fiber in the ipsilateral ICc. Acoustic and laser stimulus parameters were identical to the approach used for ICc recordings. Since animals at this stage had all completed behavioral training and assessment, the peak amplitude used for acoustic and laser stimulation was set to a suprathreshold level according to the corresponding behavioral from each mouse (60dB SPL and 12dB above the laser detection threshold, respectively).

**Behavioral testing**

Behavioral training was carried out in an acoustically transparent enclosure (8× 6× 12 inch, L× W× H) bisected into two virtual zones resting atop electrified flooring (8 pole
scrambled shocker, Coulbourn Instruments). Mouse position was tracked with a commercial PC webcam. Auditory stimuli were delivered through a free-field speaker positioned above the apparatus to provide a relatively homogenous sound field (Tucker-Davis Technologies). Mice were given at least five minutes to acclimate to the apparatus before each day of testing. Naïve mice were initially shaped to cross between zones of the chamber to terminate a foot shock (60Hz, 0.5–1mA, according to the minimally effective intensity for each mouse). With conditioned crossing behavior established, mice were then trained to associate sound (white noise, 5 s duration, 5ms cosine ramps, 70dB SPL) with foot shock initiated 5 s later. Crossing within the 5 s window was scored as a hit and the foot shock was avoided. Foot shock was initiated if the mouse failed to cross within the 5 s period (a miss) and was terminated upon crossing sides or 10 s, whichever occurred first. Once the hit rate stabilized at ≥ 70%, white noise was replaced with the narrow-band noise bursts and training continued until crossing behavior stabilized again. Psychometric functions were acquired by documenting the hit probability at different sound levels (−10 to 70dB SPL in 10dB steps) and pulse rates (60–300Hz in 60Hz steps). Stimuli were presented in a pseudo random fashion and repeated at least 15 times each. Inter-trial intervals were randomly drawn from a uniform distribution between 30 to 40 seconds. False positives were calculated as animal’s crossing probability during a 5 s window halfway through the inter-trial period. Typically, each animal performed 60 to 100 trials per day, 5 to 6 days per week.

For behavioral experiments involving detection of laser pulse trains rather than acoustic pulse trains, the implanted midbrain assembly was tethered to the laser with a patch cable. Mice were given approximately 20 minutes to acclimate to tethering before conditioned crossing behavior was initially reestablished with broadband noise stimulus. Once the hit probability was
comparable to that documented without tethering, the acoustic stimulus was replace by laser stimuli (1ms laser pulses, 60 to 300Hz in 60Hz steps) without any additional behavioral shaping. Due to the variability of sensitivity introduced by injection volume and expression level of the opsins, the range of laser intensity tested was adjusted on a case-by-case basis for each animal to generate a range of subthreshold to suprathreshold behavioral responses. In all other respects, stimulus design and task organization were identical to the acoustic version of the task.

**Histology**

Animals were deeply anesthetized with ketamine and prepared for transcardial perfusion with 4% formalin solution. The brains were extracted and post-fixed in 4% formalin at room temperature for an additional 12 hours before transfer to 30% sucrose solution. Brain sections (60μm thick) were counterstained with DAPI (Life Technologies). The position and size of the infection zone was inferred through visualization of the fluorescent label with a conventional epifluorescence microscope (Zeiss).

**Data analysis**

Firing rate adaptation was quantified by calculating the ratio of the spike count to the first pulse divided by the average spike count to all remaining pulses within the 1 s period. To quantify the temporal fidelity of sound or laser evoked activities, a template-based classifier model was used. For any given recording site, half the trials of responses to all pulse rates were used to build peristimulus time histogram (PSTH) based templates; the other half of the trials were used as test cases. Test trials were compared with the templates by calculating their cross correlation coefficients. The decoded pulse rate for a test trial was the pulse rate behind the most similar template (highest cross correlation coefficient). The decoding accuracy for all rates was calculated and averaged across recording sites. Pulse train detectability was quantified by
dividing the PSTH into 100ms bins and calculating the firing rates for each bin within the spontaneous and evoked periods on a single trial basis. For any given bin, its detectability was quantified as the rectified z-score of its spike count with respect to the baseline distribution. The difference between mean z-scores from the spontaneous and evoked periods for each trial provided the basis for calculating d’.

In the population version of this analysis, the neural detectability for a given stimulus was calculated as the highest detectability provided by any of the recorded sites from an animal.

**Statistical analyses**

All statistical analyses were performed in Matlab (Mathworks). Repeated measures ANOVAs were used to compare neural or behavioral measurements over dependent variables such as pulse rate or sound intensity in the same group of animals. When comparing measurements across different groups of animals, mixed-designed ANOVAs were used, and the main effects were reported. Multiple comparisons were corrected with the Bonferroni method.
Figure 4.3: Chronos-mediated spiking provides a superior basis for pulse rate detection in the midbrain.

(a), Firing rates from individual trials within a 2s period surrounding the onset of photostimulation in a Chronos\(^+\) mouse is plotted in 0.1s bins. (b), The distribution of single trial firing rates from each bin falling within the
spontaneous (-1 to 0s) and evoked (0 to 1s) periods, derived from the results in panel A. (c). For all three forms of stimulation, rates from both periods are then converted into z-scores and the absolute values used as rectified z-scores, where more positive values represent deviations in firing rate that could support a neural code for detection. (d), the separation between the spontaneous and evoked z-score distributions is expressed with the d’ metric to quantify the utility of a rate code for detecting sound or photostimulation.

RESULTS

Temporal variations in the sound pressure envelope provide the dominant physical cue for speech intelligibility (Shannon et al. 1995). To understand more about the accuracy and limitations of temporal coding with optogenetic stimulation of the central auditory pathways, we characterized the temporal response fidelity of midbrain neurons activated by trains of narrow-band noise (NBN) bursts or pulses of light at presentation rates ranging from 20–300Hz. In these experiments, extracellular recordings of multiunit activity in the ICc were made several weeks after the same brain area was infected with a viral construct encoding either ChR2 or Chronos (Fig. 4.1a). In response to natural acoustic stimulation delivered to the contralateral ear, ICc neurons synchronize action potential timing to NBN pulse rates as high as several hundred Hz, with gradually adapting non-synchronized responses observed at higher rates (Fig. 4.1b). In ChR2 expressing regions (ChR2+), laser stimulation induced robust synchronization at low pulse rates (<50Hz) and strongly adapting, non-synchronized responses at higher rates. By contrast, Chronos+ neurons could be entrained by a substantially wider range of photostimulation rates, with more sustained, non-synchronized activity observed even for the highest pulse rates tested (Fig. 4.1c). To quantify firing rate adaptation, we calculated the spike rate ratio between the first and the rest of the stimulus pulses for optogenetic versus acoustic stimulation. We found that
adaptation increased with pulse rate for all modes of activation (Repeated-measures ANOVA, N= 388, df= 14, F= 365.4, p= 2.9×10−6) but was more pronounced overall for ChR2 compared to NBN (Mixed-design ANOVA; main effect for stimulus type; N= 117/160, df= 1, F= 22.8, p= 2.9×10−6). The opposite relationship was noted in Chronos+ neurons, in that adaptation was significantly more pronounced for acoustic pulse trains than optical pulse trains (Mixed-design ANOVA; main effect for stimulus type; N= 111/160, df= 1, F= 56.3, p= 9.0×10−13, Fig. 4.1d).

To determine how the various firing rate profiles evoked by acoustic and optogenetic stimuli translated into a robust neural code for pulse rate, we used a PSTH-based classifier model to infer the parent stimulus rate from spike trains collected on individual trials. To the extent that each pulse rate was encoded by a distinct and reliable pattern of spiking, the classifier should correctly identify the pulse rate of the parent stimulus on an individual trial basis (Fig. 4.2a). The resultant confusion matrices computed from individual recording sites display the probability of veridical stimulus classification on the diagonal with classification errors appearing on either side. Qualitatively, all three types of activation support a robust neural code at low pulse rates, yet direct photoactivation via Chronos is the only approach that supports a robust code for higher pulse rates (Fig. 4.2b–d). Looking across all recording sites, we found a significantly higher probability of correct stimulus rate classification with Chronos compared to either ChR2 or NBN (Mixed-design ANOVA; main effect for injection type; Chronos vs. ChR2: N= 117/160, df= 1, F= 30.1, p= 9.0×10−13, Chronos vs. NBN: N= 111/160, df= 1, F= 21.3, p= 6.0×10−6 , both significant after the Bonferroni correction for 3 comparisons) but no significant difference between ChR2 and NBN (Mixed-design ANOVA; main effect for stimulus type; N= 117/160, df= 1, F= 0.44, p= 0.5, not significant after the Bonferroni correction for 3 comparisons; Fig. 4.2e).
Figure 4.4: Behavioral detection of midbrain photostimulation is similar across pulse rates and opsin types. (a), Detection abilities were tested with an avoidance paradigm where mice crossed from one side of a shuttlebox to another to avoid receiving an electric shock (top). Crossing during the 5s cue period was scored as a hit while
The previous analyses characterize the ability of midbrain neurons to discriminate between different pulse rates but leave open the more basic question of whether and how different methods of stimulation might support a more rudimentary code for simple detection of a stimulus. We addressed this point by counting the number of spikes in each 0.1 s bin for each trial (Fig. 4.3a) and comparing the overall distribution of spikes during the spontaneous and evoked periods (Fig. 4.3b). Spike counts from the spontaneous and evoked periods of the PSTH were then converted into z-scores and rectified (because any deviation from a baseline firing rate – be it an increase or decrease – could potentially drive detection), such that more positive z-scores represented a more salient neural cue for detection (Fig. 4.3c). We then computed the sensitivity index, d’, as the separation between the firing rate distributions evoked by the pulse train (hits) versus activity occurring during the pre-stimulus period (false positives) in units of standard deviations. We observed that both NBN and Chronos provided a salient cue for detection while the strong firing rate adaptation found with ChR2 stimulation interfered with a robust midbrain code for sound detection, particularly at high stimulation rates (Mixed-design ANOVA; between subjects factor; ChR2 vs. NBN: N= 117/160, df= 1, F= 44.3, p= 1.5×10−10, ChR2 vs. Chronos: N= 111/160, df= 1, F= 56.2, p= 1.5×10−12, both significant after the Bonferroni correction for 3 comparisons; Fig. 4.3d). The clear differences in neural d’ that
emerged from our midbrain neurophysiology experiments inspired a set of core hypotheses related to the perceptual salience of central auditory activation via sound versus direct photoactivation of midbrain neurons: i) behavioral detection of NBN should be robust and relatively insensitive to pulse rate; ii) behavioral detection of optogenetic stimulation should decrease at higher pulse rates; iii) Chronos should support an enhanced behavioral detection of midbrain photostimulation compared to ChR2.

To address these hypotheses, mice were trained to report detection of pulse trains delivered via acoustic stimulation or optogenetic stimulation of the midbrain. Shortly after receiving unilateral ICc injections of viral constructs or saline, mice were trained and tested in an auditory avoidance task. Specifically, NBN pulse trains of variable level and rate were presented, whereupon mice crossed sides of a shuttlebox to avoid receiving a foot shock (Fig. 4.4a). Once psychometric functions were obtained for all acoustic pulse rates (Fig. 4.4b), an optic fiber assembly was implanted into the previously injected ICc and the behavioral procedure was repeated with photostimulation in lieu of acoustic stimulation (Fig. 4.4c). Increasing stimulation amplitude was associated with a monotonic increase in behavioral detection probability (Fig. 4.4b,c). The slope of the growth function for each pulse rate could be estimated from a linear fit of the data, which provided a behavioral proxy for stimulus salience.

As predicted from the high neural d’ values for acoustic stimulation at all pulse rates (Fig. 4.3d), detection slopes were similarly steep across all NBN pulse rates. Psychometric detection slopes did not vary between injection types, but were affected by pulse rate (Mixed-design ANOVA; main effect for injection type: N= 3/3/2, df= 2, F= 1.1, p= 0.40; main effect for pulse rate: N= 3/3/2, df= 4, F= 6.6, p= 1.5×10−3, Fig. 4.4d). The d’ sensitivity index was then computed at a fixed sensation level (50dB above the detection threshold, which usually generates
hit rates around 70%). This analysis supported the conclusion that detection of acoustic pulse trains was robust for all injection types, and modestly affected by pulse rate (Mixed-design ANOVA; main effect for injection type: N= 3/3/2, df= 2, F= 0.32, p= 0.74; main effect for pulse rate: N= 3/3/2, df= 4, F= 7.0, p= $1.0 \times 10^{-3}$, Fig. 4.4e).

Upon switching the detection stimulus from sound to light we found that Chr2+ and Chronos+ mice immediately generalized across stimulation modalities (Movie S1). To our surprise, however, Chronos+ and ChR2+ mice exhibited similar detection slopes (Mixed-design ANOVA; main effect for injection type: N= 3/3/2, df= 2, F= 0.13, p= 0.73; Fig. 4.4f) and d’ values (Mixed-design ANOVA; main effect for injection type: N= 3/3/2, df= 2, F= 0.05, p= 0.83; Fig. 4.4g) that did not vary as a function of pulse rate (Mixed-design ANOVA; main effect for pulse rate; slope: N=3/3/2, df=4, F=1.4, p=0.32; d’: N=3/3/2, df=4, F=0.83, p=0.52). Detection was at chance in saline-injected mice, confirming that detection of photostimulation trains could not be ascribed to visual detection of reflected laser light (Paired t-test between hit rates and FP rates, N=10, p=0.34). Similar results were obtained when behavioral crossing latency was substituted for crossing probability (Fig. S1). Thus, behavioral findings in channelrhodopsin-expressing mice were essentially opposite to the predictions based on our ICc recordings, from which we surmised that Chronos+ mice would be more sensitive than ChR2+ and that perceptual salience be reduced at higher pulse rates.

One potential explanation for these discrepant findings stems from the possibility that robust midbrain neural codes for the discrimination (Fig. 4.2) and detection (Fig. 4.3) have been reformatted, diminished, or lost altogether upon reaching higher stations of the auditory CNS, where neural activity is more intimately linked to perception (Atiani et al. 2014; Niwa et al. 2012; Jarmamillo and Zador 2011; Buran et al. 2014; Schneider and Woolley 2013; Bendor and
Wang 2007; Bizley et al. 2013; Mesgarani et al. 2012). For example, neural activity in the auditory cortex (AC) might be organized in a manner that more closely resembles behavioral detection functions than midbrain activity. We addressed this hypothesis by recording activity from the core fields of AC (the primary auditory cortex and the anterior auditory field) in the same mice that were tested behaviorally. For the sake of comparison to the ICc data, we recorded multiunit activity evoked by acoustic pulse trains and midbrain optogenetic stimulation under anesthesia with similar stimulation and recording protocols (Fig. 4.5a).

Optic fiber tip implantation and virus injections were made at 0.35mm and 0.7mm below the dorsal surface of the inferior colliculus, respectively. These depths approximately correspond to a best frequency range of 10–22 kHz within the ICc tonotopic map, respectively (Fig. 4.5a, S2). By sampling optogenetically-evoked spiking across the AC tonotopic map, we confirmed that feedforward activation from the midbrain was greatest within zones of the AC with similar best frequencies (Two-way ANOVA; main effect for BF; N=42/45, df=4, F=3.97, p=0.54× 10−3 ; Fig. 4.5b). Aside from this topographic correspondence, the downstream effects of optogenetic stimulation measured in the AC bore little resemblance to the direct activation of midbrain neurons. Whereas acoustic stimulation evoked non-synchronized, partially adapting responses in AC, photostimulating midbrain ChR2+ (Fig. 4.5c) or Chronos+(Fig. 4.5d) ICc neurons evoked brief, non-synchronized onset responses (<50ms) followed by suppression and a rapid return to baseline (Fig. 4.5e). As such, sound-evoked responses in AC were significantly greater overall than either mode of direct midbrain activation (One-way ANOVA; NBN vs. ChR2: N=136/64, df=1, F=1540.67, p < 1× 10−196).
Figure 4.5: Auditory cortex encoding of midbrain photostimulation matches behavior detection, not ICc neural detection.

(a) AC recordings were made following the completion of behavioral testing using a laser power corresponding to a 0.7 hit probability for each pulse rate. The optical fiber tip was positioned to photostimulate a mid-frequency region of the ICc tonotopic map, which (b) activated a corresponding region of the AC tonotopic map. (c-d), Rastergrams from individual recording sites depict stimulation with 1 s trains of NBN bursts or pulsed blue light. (e), PSTHs
Figure 4.5 (Continued)

averaged across all pulse rates reveal brief, weak onset responses followed by suppression with optogenetic stimulation and comparatively robust, less adapting responses evoked by NBN. (f-g). Rectified z-scores and the resultant d’ index were calculated for each pulse rate and stimulation type as described previously for ICc recordings in Figure 4.3.

To determine if AC neural detection across pulse rates more closely resembled midbrain or behavioral detection functions, we set the laser power for each pulse rate to be 12dB above the detection threshold for each pulse rate, where hit rates were generally greater than 70%. We found that AC neural detection of optogenetic stimulation was fairly constant across pulse rates and was similarly robust for ChR2 and Chronos (but lower than for NBN), both in terms of both the z-scores separating the evoked and spontaneous periods (Fig. 4.5f) as well as the derivative d’ values (Fig. 4.5g). For NBN inputs, d’ increased significantly across pulse rates in a similar fashion to ICc recordings (Repeated measures ANOVA, N= 160, df= 4, F= 69.6, p= 0,Fig. 4.3d) and was significantly greater overall than for either mode of optogenetic stimulation (Mixed-design ANOVA; main effect for injection type; NBN vs. ChR2: N= 136/48, df= 1, F= 17.7, p= 4.1×10−5 ; NBN vs. Chronos: N= 136/72, F= 17.2, p= 4.9×10−5 ). In contrast with ICc results, neural detection of midbrain photostimulation did not significantly differ across opsin type (Mixed-design ANOVA; main effect for opsin type: N=48/72, df=1, F=0.64, p=0.42). Although there was a significant effect of pulse rate (Repeated measures ANOVA; ChR2: N= 48, df= 4, F= 4.7, p= 1.1×10−3 ; Chronos: N= 72, df= 4, F= 8.2, p= 3.0×10−5 ), the absolute change in AC detection with either type of photostimulation was modest.

To determine if AC spiking represented the temporal features of midbrain photostimulation at the population level in a manner that could not be appreciated from an
analysis of single recording sites, we applied a population based approach to calculate the z-scores of cortical neural responses, where for a given stimulus, its population neural z-score is the highest z-score out of all the recorded sites. However, even at the level of a distributed cortical population code, midbrain photostimulation was similarly poor for both opsins (Mixed-design ANOVA; main effect for opsin type: N=3/3, df=1, F=0.05, p=0.84, Fig. S2). Thus, the cortical encoding of optogenetic pulse trains did not support the second and third hypotheses that arose from our midbrain findings, but closely resembled – and likely enabled – similar observations made at the level of behavior.

Discussion

Auditory prostheses have provided over 300,000 deaf individuals with an awareness of environmental sound and an ability to comprehend speech in quiet backgrounds. In this regard, the auditory prosthesis is a triumph of modern biomedical engineering. Speech intelligibility in quiet has steadily improved among cochlear implant recipients since the device came into common use approximately 30 years ago. Average cochlear implant users who are fitted with modern electrodes and electronic processing strategies correctly comprehend approximately 58% of monosyllabic words presented in silence (Wilson and Dorman 2008). However, the inter-subject variability in speech comprehension is enormous, even after controlling for mitigating variables such as onset and duration of deafness. For example, individuals implanted at the same clinic by the same surgeon using identical hardware, exhibit speech recognition scores ranging from 0–100% (Tyler 2000). Moreover, such devices rarely allow users to extract intelligible information from complex, degraded acoustic signals such as speech in noise. Speech recognition with auditory brainstem or midbrain implants is even worse, with users reporting
35% (n=60) and 0% (n=3) improvements in comprehension, respectively, for measures of speech performance (Lim et al. 2009; Colletti et al. 2009).

The present project was motivated by a desire to explore alternative ways to provide an actionable sound percept through direct brain stimulation. We took advantage of a recently described channelrhodopsin, Chronos, that supports non-adapting, high-fidelity encoding of rapid light pulses that fall within the range of pulse rates used in auditory brainstem implant processors (SPEAK processing strategy, 250 pulses/s, Cochlear Clinical Guidance Document 2010). Direct comparison of acoustic versus optical activation in ChR2+ or Chronos+ ICc neurons demonstrated that Chronos supported a superior encoding of temporal pulse trains, in terms of firing rate adaptation (Fig. 4.1), temporal coding (Fig. 4.2) and the overall salience of neural responses at high stimulation rates (Fig. 4.3). These performance advantages were not realized when the downstream effects of midbrain photostimulation were assessed behaviorally (Fig. 4.4) or through AC recordings (Fig. 4.5).

We found that midbrain photostimulation induced weak cortical onset responses followed by firing rate suppression across a broad range of stimulation frequencies, consistent with previous descriptions of cortical response suppression upon electrical stimulation of the inferior colliculus in animals (Calixto et al. 2012; Straka et al. 2014). Thus, a spatiotemporally coherent barrage of afferent activity might be ill-suited to engage the distributed coding capacities of cortical neurons. One possible explanation lies in the observation that fast-spiking parvalbumin+ inhibitory interneurons, which are known to rapidly ‘quench’ excitatory responses in primary sensory cortex (Cardin et al. 2009; Li et al. 2013), have broader frequency tuning and lower response thresholds than putative pyramidal neurons (Li et al. 2014, but see Ref. Moore and Wehr 2013). Thus, simultaneously exciting a relatively broad region of the subcortical tonotopic
map might preferentially engage cortical inhibition, thereby preempting robust, sustained responses in cortical excitatory neurons, which are more responsive to afferent regimes characterized by spatially differentiated patterns of inputs organized on multiple, nested time scales (Ulanovsky et al. 2004; Engineer et al. 2008; David et al. 39; Zhou and Wang 2010; Rabinowitz et al. 2012).

An alternative explanation holds that cortical networks do indeed encode temporal information contained within the photoactivated midbrain efferent signals, but that the nature of the code has been reformatted from a temporal scheme that is evident in individual midbrain recording sites to a spatially distributed, rate-based scheme that can only be appreciated when documented at the level of coordinated activity an ensemble of cortical neurons. Indeed, internally or externally generated inputs to auditory cortex are processed by locally organized subnetworks of cortical neurons (Rothschild et al. 2010; Bandyopadhyay et al. 2010; Bathellier et al. 2012; Issa et al. 2014) that feature highly stereotyped spatiotemporal dynamics (Luczak et al. 2009) that generate inter-neuronal correlations, which may (Jeanne et al. 2013) or may not (Rothschild et al. 2013) augment AC coding capabilities. Although our approach did not permit the direct visualization of cellular ensembles in the cortex (Fig. S3), behavioral detection exhibited a similarly weak correspondence to midbrain encoding functions, suggesting that even if temporal features were encoded that could not be observed with our neurophysiological approach they are not being translated into behavioral detection abilities.

If temporal coding deficiencies can be attributed to the stimulation strategy and not to the readout, how could one design a central sensory prosthesis better suited to the encoding preferences of higher brain networks so as to provide a more robust percept? We suggest that the core deficiency lies in the inability to selectively activate the projection neurons of low-level
sensory brain areas with spatially differentiated inputs. Whereas a single cochlear implant channel activates a relatively homogenous population of cochleotopically restricted afferent nerve fibers, direct stimulation of the auditory brainstem or midbrain activates a heterogeneous and spatially undifferentiated population of neurons and glial cells. This deficiency is insurmountable with electrical stimulation strategies but could be addressed through further development of an optogenetics-based prosthesis.

What would be the defining features of a future optoprostheses? Ideally, the channelrhodopsin could be selectively expressed in primary output neurons at the earliest stage of central auditory processing (e.g. globular and spherical bushy cells, multipolar, and fusiform neurons (Adams 1979; Brunso-Bechtold et al. 1981)). Cell-type specific expression is trivial to realize in laboratory animals, where Cre-loxP transgenic recombination systems used in conjunction with Cre-dependent viruses ensure that transcription of viral constructs are limited to the cell type of interest (Madisen et al. 2012). An alternative approach that might be more feasible in humans would be to design viral constructs with promoter sequences that are transcribed at far higher levels in the cell type of interest than neighboring cells. In this manner, transcription of the channelrhodopsin DNA could be ‘steered’ into particular cell types within the cochlear nucleus. This approach has been used with some success in the cerebral cortex, where viral constructs packaged with the CAMKIIa promoter are expressed at far greater levels in pyramidal neurons due to natural variations in expression of this gene between different cell types (Dittgen et al. 2004). In situ hybridization studies of the cochlear nucleus or its avian homolog reveal striking differences in mRNA levels between cell types, which lends provisional support to further study of transcriptional differences between principal neurons and other cell
types of the cochlear nucleus (e.g. mGluR1α in globular bushy cells Bilak et al. 1996; VGlut2 in fusiform neurons Hackett et al. 2011c).

The challenge of creating spatially differentiated input patterns could be tackled through an engineering approach to develop bundles of low-power, miniature-scale independent illumination systems or a biological approach to engineer channelrhodopsins that are sensitive to restricted, non-overlapping wavelengths (Klapoetke et al. 2014). With this latter approach, the brainstem could be ‘seeded’ with constructs encoding distinct channelrhodopsins that were packaged in viral serotypes selected to provide highly focal infection zones. Alternatively, a single construct encoding multiple channelrhodopsins with a self-initiated transcriptional suppressor system might be used to drive stochastic expression of a single channelrhodopsin from the cassette, similar to the approach used for multi-colored fluorescent protein expression in ‘Autobow’ mice (Cai et al. 2013).

Provided a minimal level of spatiotemporally differentiated activity, ample time for adjustment, and a cellular milieu that supports an acceptable degree of plasticity, higher levels of the brain can support a remarkable constellation of perceptual operations despite a distorted or impoverished input signal. In classic examples, humans and other animals rapidly adapt to visual goggles that dramatically shift the visual field (Harris 1965) or to temporary manipulations of the outer ear that grossly distort dichotic cues that underlie sound localization (Hoffman et al. 1998). Perhaps even more impressively, postlingually deafened cochlear implant users can map the massively altered and degraded patterns of electrical stimulation onto prior neural representations of acoustic speech sounds. Further, they can rapidly recover or exceed baseline speech recognition when temporal stimulation strategies are radically changed (Wilson et al. 1991) or can achieve a fused pitch percept from bilateral implants that stimulate grossly
mismatched cochlear positions (Reiss et al. 2011). Thus, whatever the form optoprosthesis development may take, future recipients, surgeons, and scientists can take solace from the observation that the prosthetic device need not provide a perfect signal, but rather one that is just good enough to fuel the remarkable pattern recognition abilities of the brain.
CHAPTER 5

CELLULAR AND WIDEFIELD IMAGING OF SOUND FREQUENCY
ORGANIZATION IN PRIMARY AND HIGHER-ORDER FIELDS OF THE MOUSE
AUDITORY CORTEX

Sandra Romero*, Ariel Edward Hight*, Kameron K. Clayton, Jennifer Resnik, Ross S. Williamson, Kenneth E Hancock, Daniel B Polley

* denotes equal contribution

Introduction

Topographic connections between the sensory receptor epithelia and downstream brain nuclei form well in advance of sensory experience and reflect the patterning of molecular guidance cues (Rakic et al. 2009; Cramer and Gabriele 2014). In the rodent auditory system, topographic connections linking the medial geniculate body (MGB) of the thalamus and the auditory cortex (ACtx) form in the last week of embryonic development, approximately one week prior to the onset of spontaneously generated intrinsic signaling between the ear and the brain, and 2 weeks prior to ear canal opening and the onset of hearing (Gurung and Fritzsch 2004; Tritsch et al. 2007; Polley et al. 2013). Cortical maps are not organized into a single, topographic gradient, but rather as a mosaic of repeating gradients, separated from one another by “mirror reversals” in receptor epithelium mapping. These individual fields of the ACtx can exhibit specializations in functional processing that reflect different origins of thalamic inputs as well as regional variations in the source, though not precision, of local and long-range intracortical connections (Rose and Woolsey 1949a; Lee and Winer 2005; Winer et al. 2005).
Core fields of the ACtx receive their predominant thalamic input from ventral division of the MGB (MGBv), which confers well-defined frequency tuning arranged into smoothly varying tonotopic gradients (Merzenich and Brugge 1973; Andersen et al. 1980; Winer et al. 2005; Hackett 2011). Higher-order ACtx fields are innervated by non-primary divisions of the MGB and from intracortical inputs originating outside of the auditory cortex (Andersen et al. 1980; Reale and Imig 1980; Schreiner and Cynader 1984; Lee and Winer 2005; Winer et al. 2005; Hackett 2011a). Higher-order cortical areas show stronger selectivity for processing conspecific communication sounds (Schneider and Woolley 2013; Norman-Haignere et al. 2015), greater involvement in cross-modal plasticity (Lomber et al. 2010) and stronger state-dependent modulation by cognitive influences such as task demands and learning (Dong et al. 2013; Atiani et al. 2014; Elgueda et al. 2019).

The mouse is among the most popular model systems for studies of cortical sound processing and plasticity, but fundamental aspects of core and higher-order cortical field organization remain unclear. The mirror reversal in tonotopy between the primary auditory cortex (A1) and an anterior auditory field (AAF), which has been observed in dozens of species (Kaas 2011), has been questioned in the mouse where instead several groups have described a region at the border of A1 and AAF without well-defined selectivity for pure tones (Issa et al. 2014, 2016; Liu et al. 2019). The systematic mapping of preferred frequency in core fields is well-accepted at mesoscale resolution (Stiebler et al. 1997; Hackett et al. 2011a; Guo et al. 2012), but remains a point of contention at the cellular scale, with some reports describing a heterogeneous salt and pepper organization, while others describe a precise relationship at all spatial scales (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Winkowski and Kanold 2013; Issa et al. 2014; Panniello et al. 2018; Liu et al. 2019; Tischbirek et al. 2019). The secondary
field, A2, was named without having established that its primary thalamic input arose from the higher-order subdivisions of the MGB (Stiebler et al. 1997). Instead, recent evidence reports that A2 receives its primary input from the MGBv, with minimal input from the higher-order dorsal subdivision, the MGBd, raising questions about whether there are any fields in the mouse ACtx that are appropriately described as higher-order and, if so, where they are located (Ohga et al. 2018). Even the name and location of A1 is not beyond dispute, where some groups refer to it instead as a “dorsomedial field” (Tsukano, Horie, Ohga, et al. 2017).

Confusion about the basic features of mouse ACtx organization stems to no small extent from differences in methodology. Traditional microelectrode mapping cannot reveal local organization at a cellular scale and cannot easily be performed in awake animals. Imaging of intrinsic hemoglobin or flavoprotein fluorescence signals can be used to visualize the entire ACtx at once, but provide a less-direct, even lower resolution map than multiunit microelectrode recordings (Kalatsky et al. 2005; Tsukano, Horie, Ohga, et al. 2017; Tischbirek et al. 2019). 2-photon imaging of calcium signals through bulk-loaded dyes or virus mediated gene transfer can be performed in awake mice and provides direct access to a cellular signal closely linked to spiking, but does not easily provide the even, stable expression over large areas needed to image multiple fields of the ACtx (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Kato et al. 2016; Kuchibhotla et al. 2017; Francis et al. 2018; Tischbirek et al. 2019). A promising approach to resolve the disputed organization of the mouse ACtx at multiple scales comes from imaging of awake transgenic mice that express the genetically encoded calcium indicator GCaMP6 in select cell types (Issa et al. 2014, 2016; Babola et al. 2018; Panniello et al. 2018; Liu et al. 2019). Here, we performed multiscale imaging from excitatory neurons in the ACtx of Thy1-GCaMP6s.
reporter mice to delineate mesoscale map organization through widefield, epifluorescence imaging and cellular organization of frequency selectivity through 2-photon imaging.

Methods

Experimental model and subject details

Adult mice of either sex were used for all experiments in the study. All procedures were approved by the Massachusetts Eye and Ear Infirmary Animal Care and Use Committee and followed the guidelines established by the National Institute of Health for the care and use of laboratory animals. Thy1-GCaMP6s mice (Jackson labs stock number 025776) were used in a subset of auditory brainstem response (ABR) measurements. Imaging experiments were performed on male and female offspring of Thy1-GCaMP6 crossed with CBA/CaJ, a strain that retains good ABR thresholds through adulthood (Zheng et al. 1999). Mice were housed in group cages until cranial window implantation, at which point they were housed individually. Mice were maintained in a 12h light/dark cycle, with food and water available ad libitum. Mice were between 5-7 weeks at the time of cranial window surgery and were no older than 20 weeks by the time the last imaging session was completed.

Auditory brainstem response measurements.

Mice were anesthetized with ketamine and xylazine (100/10 mg/kg for ketamine/xylazine, respectively, with boosters of 50 mg/kg ketamine given as needed). Core body temperature was maintained at 36.5° with a homeothermic blanket system. Acoustic stimuli were presented via in-ear acoustic assemblies consisting of two miniature dynamic earphones (CUI CDMG15008–03A) and an electret condenser microphone (Knowles FG-23339-PO7) coupled to a probe tube. Stimuli were calibrated in the ear canal in each mouse before recording.
ABR stimuli were tone bursts (4-64 kHz in 0.5 octave increments), 5 ms duration with a 0.5 ms rise-fall time delivered at 27 Hz, and alternated in polarity. Intensity was incremented in 5 dB steps, from 20-80 dB SPL, or as high as 100 dB SPL in cases with elevated thresholds. ABRs were measured with subdermal needle electrodes positioned beneath both pinna (+ and -) and the base of the tail (ground). Responses were amplified (gain = 10,000), filtered (0.3–3 kHz), and averaged (1024 repeats per level). ABR threshold was defined as the lowest stimulus level at which a repeatable wave 1 could be identified. ABR Recordings were made from Thy1-GCaMP6s mice at three ages: 7-8 weeks (9 ears, 5 mice), 14 weeks (10 ears, 5 mice) and 20 weeks (8 ears, 4 mice). ABR recordings were made from Thy1-GCaMP6s × CBA mice at the same ages: 7-8 weeks (11 ears, 6 mice), 14 weeks (6 ears, 3 mice) and 20 weeks (6 ears, 3 mice).
Figure 5.1: A transgenic GCaMP6s reporter mouse that retains good hearing into adulthood.

(A) Auditory brainstem response waveforms elicited with 45 kHz tone bursts at various sound levels in a representative 7-week old Thy1-GCaMP6s mouse (gray) or the F1 offspring of Thy1-GCaMP6s crossed with CBA (black). Blue arrows denote wave 1. Red arrows point towards wave 1 threshold sound level. (B) Mean ± SEM ABR thresholds measured at various postnatal ages from Thy1-GCaMP6s (n = 27 ears from 14 mice) and Thy1-GCaMP6s × CBA mice (n = 23 ears from 12 mice). Bottom right sub-panel presents median and 90% confidence intervals for high-frequency thresholds (32-64 kHz). Asterisks denote p < 0.05 with Wilcoxon Rank-Sum test. (C) Confocal image of ACtx GCaMP6s labeling across cortical layers in the Thy1-GCaMP6s × CBA mouse. Scale bar: 100 microns (D) Rendering of custom head fixation hardware and placement of the chronic cranial window on mouse skull. Inset: Cranial landmarks used to position the window atop the ACtx include the lambdoid suture (LS), temporal ridge (TR) and zygomatic arch extension (ZAE).

Preparation for chronic imaging.

Glass cover slips were first etched in piranha solution (H2O2 mixed with H2SO4 in a 3:1 ratio) and stored in 70% ethanol. A 4mm diameter cover slip was centered and affixed to a pair of 3mm cover slips (#1 thickness, Warner Instruments) using a transparent, UV-cured adhesive (Norland Products). Windows were stored in double deionized water and rinsed with sterile saline before use. On the day of surgery, animals were anesthetized with isoflurane in oxygen (5% induction; 1.5-2% maintenance). After removing the periosteum from the dorsal surface of the skull, an etchant (C&B Metabond) was applied for 30 sec to create a better adhesive surface. Custom stainless-steel head fixation hardware (iMaterialise) was bonded to the dorsal surface of the skull with dental cement (C&B Metabond) mixed with India ink. A 3mm circular craniotomy was made atop the ACtx with the combination of a scalpel and a pneumatic dental drill with diamond burr (head diameter 1/10 mm, NeoDiamond – Microcopy Dental). The coverslip was then lowered into place using a 3-D manipulator and bonded to the surrounding regions of the skull to create a hermetic seal. Post-operative injections of Buprenex (0.05 mg/kg) and
Meloxicam (0.1 mg/kg) were administered and the mice were allowed to recover in a warmed chamber. Imaging began 5-7 days later.

**In vivo calcium imaging.**

Widefield epifluorescence images were acquired with a tandem-lens microscope (THT-microscope, SciMedia) configured with low-magnification, high-numerical aperture lenses (PLAN APO, Leica, 2× and 1× for the objective and condensing lenses, respectively). Blue illumination was provided by a light-emitting diode (465 nm, LEX2-LZ4, SciMedia). Green fluorescence passed through a filter cube and was captured at 20Hz with a sCMOS camera (Zyla 4.2, Andor Technology). Two-photon excitation was provided by a Mai-Tai eHP DS Ti:Sapphire-pulsed laser tuned to 940 nm (Spectra-Physics). The beam spot size was controlled with variable expander optics (Thorlabs) and the intensity was adjusted with a variable attenuator (Thorlabs) and Pockels cells (Conoptics). Imaging was performed with a 16×/0.8NA water-immersion objective (Nikon) from a 512 x 512 pixel field of view at 30Hz with a Bergamo II Galvo-Resonant 8 kHz scanning microscope (Thorlabs). Scanning software (Thorlabs) was synchronized to the stimulus generation hardware (National Instruments) with digital pulse trains. Widefield and 2-photon microscopes were rotated by 50-60 degrees off the vertical axis to obtain images from the lateral aspect of the mouse cortex while the animal was maintained in an upright head position. Imaging was performed in light-tight, sound attenuating chambers (N=12 mice for widefield imaging and N=4 mice for 2-photon imaging). Animals were monitored during the experiment with modified cameras (PlayStation Eye, Sony) coupled to infrared light sources. For widefield imaging, the focal plane was set to be approximately 200 μm below the pial surface. For 2-photon imaging, the imaging depth ranged from 175-225 μm below the pial surface, in layer 2/3.
Auditory stimulation for imaging experiments

Stimuli were generated with a 24-bit digital-to-analog converter (National Instruments model PXI-4461) using scripts programmed in MATLAB (MathWorks) and LabVIEW (National Instruments). Stimuli were presented via free-field tweeters positioned 10 cm (2-photon system) or 25 cm (widefield system) from the left (contralateral) ear. Free-field stimuli were calibrated before recording using a wide-band ultrasonic acoustic sensor (Knowles Acoustics, model SPM0204UD5). Pure tones were pseudorandomly presented at variable frequencies (4-64 kHz in 0.5 octave steps) and intensities (0-70 dB SPL in 10 dB steps) such that each of the 72 unique frequency-intensity combinations were presented 20 times each. Tone duration was 50 ms. Trial length was either 3s (2-photon imaging) or 3.5s (widefield imaging).
Figure 5.2: Anatomical landmarks for tonotopically organized fields in the mouse auditory cortex.

(A) Tandem lens macroscope for widefield GCaMP6s imaging. (B) Top: Raw calcium signals from a typical pixel are de-trended with a 10s moving average (gray dashed line). Bottom: Fractional change in fluorescence is computed relative to the moving average (F₀). Red lines denote timing of individual tone bursts. (C) Mean ± SEM fractional change in fluorescence from 20 repetitions of a 60 dB SPL tone before and after spatial and temporal filtering (light and dark green, respectively). (D) Histogram of fractional change values from all temporal baseline periods (gray) and from the peak amplitude of the tone-evoked responses at four frequencies. Dashed black line denotes 2 SD above the mean of the baseline values used to define tone threshold for each pixel. (E) Frequency response area of peak responses expressed as z-score from the baseline distribution. Best frequency (BF) is computed from the frequency response function derived from sound levels at threshold to threshold + 20 dB. (F) Each pixel is assigned an opacity and hue to denote the response amplitude and tone frequency corresponding to the BF, respectively. In this example case, the medial, lateral, caudal and rostral edges of the tonotopically organized areas were marked.
Figure 5.2 (Continued)

with Di-I after imaging. (G) Coronal sections of the four Di-I placements from shown in F (right) and a grayscale image from a generic mouse brain database to show approximate anatomical landmarks (left). Values express approximate distance from Bregma. Asterisk denotes rhinal sulcus. Red arrows denote center of Di-I expression. Data from panels B-E and panels F-G come from two different mice.

Image processing – Widefield

Raw data was first downsampled from the native 1200 × 1200 pixel resolution to 256 × 256 pixels. Slow drifts in the fluorescence signal were removed from the measurement by concatenating all frames for an individual imaging session and computing a temporal baseline (F0) for each pixel from the linear fit of a 10s sliding window incremented in 5s steps (Chronux toolbox, Matlab). The fractional change in fluorescence was defined for each frame as a percent change in signal from the temporally detrended average signals as $(\frac{\Delta F}{F0}) \times 100$. Individual trials were then averaged across the 20 repetitions and temporally filtered with a GCaMP6s impulse response deconvolution kernel (Chen et al. 2013). Fluorescence data were spatially denoised with Lucy-Richardson deconvolution using a Gaussian filter (470 µm width). For each stimulus, the temporal peak in the sound-evoked response period was defined independently for each pixel as the frame with the maximum percent change within a 0.75s period following stimulus onset, which was then averaged with the immediately preceding and following frame. For each pixel, baseline activity levels were defined by creating a histogram of percent change amplitudes during the 0.5s pre-stimulus period (25 frames × 72 stimuli). The response amplitude for each tone/level combination was then expressed in units of standard deviations (z-score) relative to the distribution of baseline activity levels.

Image processing – two-photon
Imaging data were processed with Suite2p, a publicly available software package that provides a complete pipeline for processing calcium-dependent fluorescence signals collected with 2-photon microscopes (Pachitariu et al. 2016; Stringer and Pachitariu 2019). Briefly, fluorescence data were collected at 2x digital magnification and processed in four stages:

*Frame registration:* Brain movement artifacts are removed through a phase correlation process that estimates the XY offset values that bring all frames of the calcium movie into register. Suite2p emphasizes correcting for movement artifacts at high spatial frequencies by first applying spatial whitening before computing a cross-correlation map. A non-rigid method is then used for phase correction that divides the movie into independent blocks and computes the optimal XY offset for each discrete segment before applying the interpolated pixel shift function to the original image.

*Detecting regions of interest (ROI):* Suite2p then identifies candidate cellular ROIs using a generative model with three key terms: i) a model of ROI activity, ii) a set of spatially-localized basis functions to model a neuropil signal that varies more gradually across space, and iii) Gaussian measurement noise. Fitting of this model to data involves repeatedly iterating stages of ROI detection, activity extraction, and subsequent pixel re-assignment.

*Signal extraction and spike deconvolution:* Suite2p then extracts a single fluorescence signal for each ROI by modelling the uncorrected fluorescence as the sum of three terms: i) a somatic signal due to an underlying spike train, ii) a neuropil trace scaled by an ROI-specific coefficient, and iii) Gaussian noise (Stringer and Pachitariu 2019). The uncorrected fluorescence is first extracted by averaging all signals within each ROI. The neuropil trace is then computed as the average signal within an annular ring surrounding each ROI. The neuropil component is different from those identified during ROI detection, which implicitly uses pixels inside ROIs,
and are not scaled by a contamination factor. Neuropil scaling coefficients and somatic fluorescence are then simultaneously estimated using an unconstrained non-negative deconvolution, using exponential kernels.

**Cellular identification:** With a fluorescence trace assigned to each identified ROI, the final stage in the Suite2p pipeline involves identifying the subset of ROIs that correspond to neural somata. Suite2p utilizes a semi-automated approach by first labelling ROIs as cells or non-cells based on various activity-dependent statistics, before a final manual curation step.

**Response amplitude calculation:** For the sake of direct comparison to widefield imaging, we computed the response amplitude of the 2-photon signal before and after neuropil correction using the same approach of identifying the peak response period for each stimulus and expressing this value as a z-score relative to the distribution of pre-stimulus baseline values.

**Registering images across sessions**

To compare separate widefield imaging sessions from the same mouse, we first obtained images of the vasculature from the mean of the raw image stack. An Affine transformation matrix (‘imregtform’ Matlab function) was then computed between any pair of imaging sessions. The optimal Affine transformation matrix to align two images was identified using gradient descent to minimize the mean squares difference between the two images, within a maximum of 10,000 iterations.

To register the 2-photon and widefield images, an image of the surface vasculature was first obtained through the 16x objective using a CCD camera under epifluorescent illumination. We identified a set of five vascular landmarks contained in both the reference image collected on the tandem-lens widefield microscope and the target image collected through the 16x objective and the CCD camera on the 2-photon microscope. Pairs of points from the images were used to
compute an affine transformation matrix and optimally align surface vasculature landmarks collected with the two imaging systems. The transformation matrix was then applied to the image acquired from the galvo-resonant scanner under 2-photon excitation. All 2-photon imaging sessions for a given mouse were registered to a single reference widefield imaging session.

*Data analysis.*

Except for the specific analysis of map changes over time, all analyses were performed only on the first widefield imaging session from each mouse. This way, no single mouse contributed more data than any other and there was no bias in selecting any particular type of imaging data.

*Response threshold estimation:* The minimum response threshold was estimated independently for each individual pixel in the widefield image or cell in the 2-photon image. Threshold was operationally defined as the lowest sound intensity for which the response to two adjacent tone frequencies were at least two standard deviations above the distribution of pre-stimulus baseline values.

*Best frequency (BF) estimation:* Frequency response functions were obtained by averaging the response at threshold, 10 dB above threshold and 20 dB above threshold. The BF was defined from the weighted sum of the responses for each of the test frequencies on an octave-based scale. Only pixels with BF response amplitudes with z-score values $\geq 2$ were used for subsequent analyses.

*Frequency tuning bandwidth:* Tuning widths for each pixel was determined from the range of frequencies with response amplitudes z-scores $\geq 2$ at 10dB above threshold.
Strength of tonotopy: For each pair of pixels, i and j, located at cortical positions pi and pj, respectively, a BF gradient vector was defined as the BF at site i minus the BF at site j, normalized by the Euclidean distance between pi and pj, all multiplied by a unit vector in the direction from pi to pj as:

\[
\overrightarrow{g_{ij}} = \frac{\log_2(BF_i) - \log_2(BF_j)}{\|p_i p_j\|} \times \frac{\overrightarrow{p_i p_j}}{\|\overrightarrow{p_i p_j}\|}
\]

The resulting gradient, \(\overrightarrow{g_{ij}}\), points from the pixel with the lower BF to the site with the higher BF and has a length proportional to the size of the change in BF normalized by the physical separation of the pixels. For each pixel \(i\), a “tonotopic vector” was defined as the vector average of all the gradients between it and all the other pixels in the same field as:

\[
\overrightarrow{v_i} = \frac{\sum_{j \in \text{Field}(i)} \overrightarrow{g_{ij}}}{N_{\text{Field}(i)}(x)}
\]

where \(\text{Field}(i)\) is the collection of pixels that belong to the same cortical field as pixel \(i\), and \(N_{\text{Field}(i)}(x)\) is the number of pixels in this field. The vector strength was calculated for each auditory field and defined as the magnitude of the vector average of all the tonotopic vectors that belong to a given cortical field:

\[
R_{\text{Field}(x)} = \frac{\sum_{i \in \text{Field}(x)} \overrightarrow{v_i}}{N_{\text{Field}(x)}}
\]
where Field(x) corresponds to all the pixels that belong to the auditory field x, NField(x) is the number of pixels in this field, and R_{Field(x)} is the vector strength of the given field.

**Similarity index:** Modules with similar frequency tuning bandwidth or response threshold were identified by thresholding the BW10 and threshold maps at the highest and lowest quartile values and identifying the regional maxima or minima of 4x4 connected neighborhood of pixels with a minimum distance to another peak of 0.25 mm. Radial vectors were drawn from each and a Similarity Index (SI) between the center and the pixels falling along each vector was computed as:

\[
SI_x = 1 - \frac{|Value_x - Value_{Seed}|}{maxD}
\]

where ValuePoint_x is the value of the response property (BW10 or ThrBF) at a pixel falling on the radial vector, ValueSeed is the value of the response property (BW10 or ThrBF) at the seed pixel; and maxD is the maximal possible difference in the response property across the map. To compute the similarity that would occur by chance, the same procedure was repeated after shuffling the positions of the pixels 10,000 times. Module size was determined as the radial distance at which the actual SI values intersected with the mean of the shuffled SI values.

**Strength of frequency tuning:** To determine the strength of frequency tuning for each pixel (widefield) or cell (2-photon), we first identified the frequency/level combination from the entire frequency response area (FRA) with the highest response amplitude. We then determined the response amplitude for the adjacent frequencies and levels and calculated the average response amplitude from these five points. Of the remaining 67 frequency-level combinations, we selected five points at random and calculated the sensitivity index, d-prime (d’), to reflect the difference between the response amplitudes near the preferred stimulus versus stimuli selected at
random. The process of selecting five random points was repeated 1000 times and the average d’ was operationally defined as the tuning quality.

*Parcellation of auditory fields:* We adapted the standard approach of defining the boundary between two adjacent fields according to reversals or abrupt shifts in the mapping of the receptor epithelia onto the cortical surface. We first identified the center of the four low-frequency points in A1, AAF, SRAF and VPAF (BF < 16 kHz). From each of these four low-frequency hubs, a set of 1440 radial vectors were drawn, at angles ranging from 0 to 360 degrees (step size = 0.25°). The mean BFs were calculated along each radial vector (± 1°). BFs were projected along each radial axis and fit to a smoothing spline ('fit' Matlab function). The reversal was defined as the point at which the first maxima was detected along the smoothed profile. If a reversal was not detected (for example at the map edges), the field boundary was drawn at the point where pixels were no longer sound-responsive according to the criteria above. To identify the boundary of the dorsal posterior field (DP), we computed the local BF gradients within A1 and created an XY map of the vector angle. The boundary between A1 and DP was aligned with the spatial shift in the BF vector phase map.

*Widefield versus 2-photon frequency tuning:* To relate the frequency tuning preference for individual neurons measured during 2-photon imaging to the underlying frequency selectivity in the widefield map, we first re-scaled the downsampled 256x256 widefield pixel map back to the native 1200x1200 pixel map. We then identified the individual pixels that correspond to the area of the neural ROI identified in Suite2p. The difference in the BF between the somatic ROI and the matching pixels of the widefield map was calculated as:

\[ BF_{diff} = \|BF_{2p} - BF_{WF}\| \]
Where $BF_{2n}$ is the BF from the 2-photon session and the $BF_{WF_n}$ is the mean BF from the corresponding widefield ROI. The BF from the somatic ROI in the 2-photon session was calculated before and after the neuropil correction was applied.

**Local BF heterogeneity:** Variation in local BF tuning was measured from neuropil-corrected 2-photon imaging data. Within a given field of view, all somatic ROIs were identified within a 50 μm radius of the reference cell. Provided that a minimum of 5 cells were identified within this area, the median BF was computed across all cells within this local neighborhood. The absolute value of the BF difference for each cell versus the neighborhood median was calculated before repeating the process with a different reference neuron. The interquartile range of this BF distribution was operationally defined as the local BF heterogeneity.

**Global tonotopic organization:** To quantify the precision of global tonotopic organization from heterogeneous local cellular frequency tuning, we projected a radial vector from the center of the low-frequency hub to the high-frequency end of the tonotopic gradient. Position along the vector was normalized from 0-1 (corresponding to the low- and high-frequency extremes of the BF vector respectively). We then identified the global tonotopic vector that passed most directly through the center of each individual field of view from a single 2-photon imaging session. Each cell was then projected onto the nearest point of this global tonotopic vector. The Pearson correlation was defined by the BF of each neuropil-corrected cell and its position on the global tonotopic vector. Confidence intervals for the local BF heterogeneity and global tonotopic correlation coefficient were calculated by bootstrapping (10,000 iterations).
Figure 5.3: Data-driven parcellation of auditory cortical fields.

(A) BF map from an example mouse. (B) Field boundaries can be established qualitatively, based on reversals or shifts in BF. Field outlines differ from the original characterization of mouse ACtx fields due to the imaging of more lateral brain areas, but boundary positions vary widely depending on “lumper” vs “splitter” biases. (C) BF changes along four vectors stretching out from the low-frequency hub in A1. (D) Boundaries (vertical lines) are placed at reversals or steep drops in signal amplitude. (E) Tonotopic vector map of BFs with boundaries established around each low frequency hub. Inset: An abrupt shift in the phase of local BF vectors is used to demarcate the boundary of the dorsal-posterior field. Divergent tonotopy at the boundary of A1 and DP can be appreciated from the progressive shift in BF vector angles with distance from the DP-A1 border. (F) Histogram of mean A1 and DP vector angles within 0.13 mm of the boundary for each individual case. (G) Resultant parcellation of five fields: dorsal posterior (DP), primary auditory cortex (A1), anterior auditory field (AAF), ventral posterior auditory field (VPAF) and the
suprarhinal auditory field (SRAF). Maps of the five auditory fields are shown for the same example mouse (left) and the remaining eleven mice used in this study.

**Histology**

A subset of mice (N=3) were anesthetized with isoflurane (5% in oxygen induction, 1.5% maintenance) after the final imaging session and the cranial window was removed. Points along the medial, lateral, rostral and caudal edges of the tonotopically organized areas were identified relative to surface vascular landmarks. A silicon probe (NeuroNexus) was mounted on a 3-D positioner and was dipped in (for 10s) and out (for 10s) ten times into red fluorescent dye (3 mg Di-I [1,1’-dioctadecyl-3,3,3’,3’-tetramethylindocarbocyanine perchlorate] per 100 µL acetone; Sigma-Aldrich, St. Louis, MO). The probe was inserted several hundred microns into the cortex at each of the four designated points and left in place for 30 minutes, re-applying the Di-I in between each placement. Following the last insertion, animals were prepared for transcardial perfusion with 4% paraformaldehyde in 0.01M phosphate buffered saline. Brains were extracted and stored in 4% paraformaldehyde for 12 hours before transferring to cryoprotectant (30% sucrose) for 48 hours. Sections (40 µm) were cut using a cryostat (Leica CM3050S), mounted on glass slides, coverslipped (Vectashield with DAPI), and imaged with an epifluorescence microscope (Leica).

**Statistical analysis**

All statistical analysis was performed with Matlab (Mathworks). Descriptive statistics are reported as mean ± SEM unless otherwise indicated. In cases where the same data sample was used for multiple comparisons, we used the Holm-Bonferroni correction to adjust for the increased probability of Type-I error. Non-parametric statistical tests were used in cases where data samples did not meet the assumptions of parametric statistical tests. The relationship
between FRA d’ and tuning heterogeneity was quantified with a permutation test iterated 500,000 times per auditory field. For any given iteration, each cell was re-assigned a d’ value at random, ensuring that the sample of d’ values was equivalent to the true distribution of d’ values. The outcome measures of interest, either the BF heterogeneity interquartile range or the global tonotopic correlation were then recomputed using the permuted data set and the linear relationship between the outcome measures of interest and the permuted d’ values were quantified for each iteration of the permuted and actual datasets with the Pearson R correlation coefficient. For plotting purposes, confidence intervals were computed as the standard error of the actual and permuted distributions. Statistical significance was established by determining the proportion of sampled permutations that exceeded the Pearson R of the true dataset. Statistical significance was defined as P < 0.05 for all tests.

Results

A transgenic mouse model for widefield calcium imaging that retains good hearing

We began our imaging studies in Thy1-GCaMP6s mice, which were developed for 2-photon imaging of the ultrasensitive genetically encoded calcium indicator GCaMP6s in cortical pyramidal neurons (Chen, Cichon, et al. 2012; Dana et al. 2014). We observed that these mice develop a substantial high-frequency hearing loss as early as 7 weeks postnatal, as indexed by wave 1 of the auditory brainstem response (ABR) (Fig. 5.1A-B). To circumvent this problem, we crossed the Thy1-GCaMP6s line to the CBA/CaJ line, which retains excellent hearing into adulthood, and confirmed that high-frequency ABR thresholds in Thy1-GCaMP6s × CBA mice were lower than the Thy1-GCaMP6s at all ages (Wilcoxon Rank-Sum, p < 0.0005 for all ages) and did not change through 20 weeks postnatal in the crossed strain (Kruskal-Wallis, Thy1-
GCaMP6s × CBA, p = 0.38; Thy1-GCaMP6s, p < 0.005). GCaMP expression in the Thy1-GCaMP6s × CBA mouse was observed in the cytoplasm and neuropil of neurons throughout the cortical column (Fig. 5.11C). Once we were satisfied with the peripheral hearing status and cortical expression levels in Thy1-GCaMP6s × CBA mice, we manufactured custom head restraint hardware that left the ears unobstructed and adapted a surgical approach to implant a glass coverslip over the lateral areas of the skull so that the full extent of the auditory fields could be visualized (Fig. 5.1D) (Goldey et al. 2014).

We performed widefield imaging in awake, head-restrained Thy1-GCaMP6s × CBA mice using a tandem lens epifluorescence microscope that provided a large (5.6mm x 5.6mm) field of view (Fig. 5.2A). We focused the microscope 0.2mm below the pial surface to de-emphasize contributions from surface blood vessels. However, widefield calcium signals collected with epifluorescence microscopes are an amalgam of many cortical layers, not just layer 2/3. Sound-evoked Ca$^{2+}$ signals were temporally de-trended (Fig. 5.2B), filtered (Fig. 5.2C) and pure tone response thresholds were independently calculated for each pixel (Fig. 5.2D). Frequency response areas (FRA) were calculated for each pixel and a best frequency (BF) was computed from the weighted sum of responses at levels ranging from threshold to 20 dB above threshold (Fig 5.2E). The ACtx was defined from the contiguous region of tone-responsive pixels, revealing a well-defined pattern of tonotopically organized gradients of BF that closely resembled prior microelectrode mapping and widefield calcium imaging datasets (Guo et al. 2012; Issa et al. 2014; Joachimsthaler et al. 2014; Liu et al. 2019).

In a subset of mice, the cranial window was removed at the conclusion of imaging and an electrode coated in Di-I was inserted into the most caudal, dorsal, rostral and lateral edge of the tonotopically organized map (Fig. 5.2F). Di-I labeling in ACtx areas with strong tonotopic
organization covered a larger area than might be expected from widely used mouse brain atlases. For instance, the lateral edge of the tonotopic zone lies just above the rhinal fissure in a region labeled as Ectorhinal cortex or temporal association cortex (TeA) in the Paxinos and Allen Brain atlases (Fig. 5.2G, second row). Further, the most medial edge corresponds to an area labeled secondary auditory cortex, dorsal area (AuD; Fig. 5.2G, third row).

Parceling the fields of mouse auditory cortex

Our widefield imaging data confirm the known arrangement of low-frequency hubs at the caudal and rostral edges of the ACtx with tonotopic gradients that branch out and collide with one another to form boundaries between individual fields (Fig. 5.3A). We did not observe a tone-insensitive region at the border of A1 and AAF (Issa et al. 2014, 2016; Liu et al. 2019). Although individual pixel thresholds and response amplitudes can be weaker in this area, we observed a clear low-high-low BF gradient across A1 and AAF, as would be expected from the mirror reversal between core fields documented in over 20 species (Kaas 2011). The tonotopic gradient of A1 was more akin to a fan radiating from its caudal low-frequency hub, than a single linear tonotopic gradient. In the rat, the ventral limb of this gradient is called the ventral auditory field, based on careful characterization of specialized selectivity for low-intensity sounds and separate anatomical inputs from the medial tonotopic limb (Wu et al. 2006; Polley et al. 2007; Storace et al. 2010, 2011). A few new features were apparent that had gone unnoticed in prior microelectrode mapping studies, presumably on account of their very lateral position (Fig. 5.3B). We noted a second low-frequency area at the caudal-lateral extreme of the ACtx map that appeared homologous to an area described in rat imaging experiments as the ventral posterior auditory field (VPAF) (Kalatsky et al. 2005). In addition, we noted a high-frequency area at the lateral extreme of A2 that was more extensive than noted in previous publications.
The traditional approach for parceling cortical fields is subjective, where bespoke boundaries are drawn at points of reversals or abrupt shifts in the tonotopic gradient. This process is not straightforward in very small cortices, where “separate” gradients can measure less than 100 microns. In these cases, demarcating one field as separate from another can reflect the psychological disposition of the experimenter as a “lumper” or a “splitter”, as much as reflecting any degree of biological ground truth (Fig. 5.3B). These observations motivated us to develop an objective approach for parceling the mouse cortex into separate fields. We projected 360 radial vectors emanating from the center of each of four low-frequency hubs located in A1, AAF, VPAF and A2. Looking at four example vectors from the A1 low-frequency zone, we plotted the BF for each pixel along the vector projection and noted points where the BF reversed or the signal dropped below threshold as boundary points (Fig. 5.3C-D). We then computed a vector map onto each of the four fields, where the length and orientation of the arrows reflect the magnitude and direction of the local BF gradients (Fig. 5.3E). The clear bifurcation of the BF gradient phase along the medial edge of A1 was used to demarcate the boundary of the dorsal posterior field (DP; Fig. 5.3E, inset), where local BF vectors became increasingly divergent with increasing distance from the A1-DP boundary (Fig. 5.3E, polar plot). Non-overlapping distributions of mean vector angle calculated near the A1-DP boundary (± 0.13mm of the boundary) in each individual mouse confirmed the tonotopic separation of A1 and DP (Fig. 5.3F).

This approach settles on a parcellation scheme intermediate to boundaries drawn with a lumping or splitting bias. These data support the position and orientation of A1, AAF and DP, as described in the original mouse electrode mapping study (Stiebler et al. 1997) and confirmed by many subsequent studies (Linden et al. 2003; Hackett et al. 2011a; Guo et al. 2012; Issa et al.
2014; Joachimsthaler et al. 2014; Shepard et al. 2015). As we have argued previously, the ultrasonic field (UF) is a misnomer as there is no reason to suggest this area is anything apart from the continuous high frequency gradient of A1, as confirmed from the continuum of BF changes from the central nucleus of the mouse inferior colliculus (Garcia-Lazaro et al. 2015). We further argue that A2 was incorrectly identified in the seminal mouse ACtx mapping study as a homologue to the secondary auditory field found in cats (Reale and Imig 1980; Schreiner and Cynader 1984). Physiologically guided iontophoretic injections of retrograde tracers identify MGBv as the predominant source of thalamic input to A2, not higher-order thalamic subdivisions ((Ohga et al. 2018), but see (Ji et al. 2016)). Further, A2 units show vigorous responses to pure tone bursts with receptive fields organized into a coarse tonotopic gradient (Guo et al. 2012; Issa et al. 2014). Absent compelling evidence that A2 receives its predominant source of input from the higher-order thalamus or exhibits any functional feature consistent with a higher-order cortical area, it seemed most prudent to name it based on its anatomical position. In rats, the most closely-related evolutionary model system to mice, the auditory field located just medial to the rhinal sulcus is widely referred to as the suprarhinal auditory field (SRAF) (Polley et al. 2007). We now refer to this field as SRAF in mice as well. Although the size and position of each auditory field varies from mouse to mouse, the overall gestalt is preserved in all cases studies here (Fig. 5.3G).
Figure 5.4: Maps and modules in the auditory cortex.

(A) BF map from another example mouse where superimposed arrows denote the direction and strength of local BF vectors for each field. (B) Distribution of individual BF vectors shown in A grouped according to field. Black lines indicate the average vectors. (C) Mean (black bars) and individual mouse (gray circles) tonotopic vector strength from the actual BF maps derived from sound levels near threshold (solid bars), at 70 dB SPL (open bars) and shuffled maps (gray bars). Lines indicate statistically significant differences with Wilcoxon Sign Rank tests after correcting for multiple comparisons. Each point in the BF map can alternatively be color coded for frequency tuning bandwidth (D) or minimum response threshold (E). (F) Mean ± SEM BW10 (top) and threshold (bottom) module
size from regions identified with asterisks in D and E are determined by computing similarity for all pixels relative to center of identified region in the actual and spatially shuffled (gray) maps. Module boundary is marked where actual similarity first overlaps with the similarity that would occur by chance (vertical dashed lines). (G) Radial distance for BW10 and threshold modules averaged across all cases (bars) or shown for each individual module (gray circles). (H) BW10 for each field averaged across all mice (blue bar) or shown for all individual mice (gray circles).

**Analysis of maps and modules in the mouse auditory cortex**

BF is the only response feature that is mapped along the extent of each field as a smoothly varying gradient (Fig. 5.4A). We quantified the strength of tonotopy by plotting the local BF phase vectors within each field from a single mouse (Fig. 5.4B, thin gray lines) and calculating the vector sum (Fig. 5.4B, thick black line). Vector strength is derived from the length of the vector sum and reflects the consistency and strength of the local BF gradients. We computed the tonotopic vector strength from a single imaging session in each mouse (N=12) and compared differences across the five fields. To estimate the tonotopic strength that would occur by chance, the BF assignment for each pixel within a field was randomized before calculating the vector strength. This process was repeated 10,000 times and the results were averaged. We observed that the tonotopically organized vector strength from the actual maps was significantly greater than the shuffled maps for all five cortical fields (Wilcoxon signed-rank tests, p < 0.005 for all fields; Fig. 5.4C). Although all fields of the mouse cortex were tonotopically organized, the strength varied between fields (listed in order of strongest to weakest: A1, AAF, SRAF, VPAF and DP; ANOVA, F = 10.49, p < 0.000001). We found that the strength of tonotopic organization in A1 was significantly greater than SRAF, VPAF and DP, but was not significantly different than AAF (post-hoc pairwise comparisons, p < 0.005 after Holm-Bonferroni correction.
for multiple comparison; A1 vs AAF, p = 0.33), in agreement with our prior electrophysiological mapping study (Guo et al. 2012). We confirmed that the strength of tonotopy in A1 is significantly reduced when BF is calculated at a single, suprathreshold level (70 dB SPL) rather than near threshold (Fig. 5.4C, post-hoc pairwise comparisons, p = 0.04 after Holm-Bonferroni correction for multiple comparison). Differences between near-threshold and suprathreshold tonotopy trended in the same direction for other fields but were not statistically significant, probably on account of being underpowered for multiple comparisons (post-hoc pairwise comparisons, p > 0.09; Fig. 5.4C solid versus open bars).

Other than BF, each of the pixels in the mouse ACtx can also be assigned a value based on tone-evoked response features such as tuning bandwidth or response threshold. As described in cats and rats, the ACtx features circumscribed modules with broad pure tone selectivity (Fig. 5.4D) or high response thresholds (Fig. 5.4E) (Schreiner and Mendelson 1990; Recanzone et al. 1999; Cheung et al. 2001; Read et al. 2001; Polley et al. 2007). A seed was positioned at the center of each individual bandwidth or threshold module and calculated the similarity of the corresponding response feature for all pixels along radial vectors fanning out in 360° from the seed. The mean similarity index across all vectors was then plotted as a function of radial distance and compared to the similarity that would occur by chance in maps where the pixels have been spatially scrambled (Fig. 5.4F, colored versus gray lines). The radial length of individual tuning modules was approximately 0.4mm for both BW10 and threshold (unpaired t-test, t = -1.08, p = 0.83; Fig. 5.4G), indicating that a few separate modules containing regions with homogenously narrow or broad frequency tuning could fall within the boundaries of a single cortical field. In terms of overall differences between fields, we confirmed prior reports that frequency tuning bandwidth was greater overall in SRAF than in A1, AAF and DP, but not
VPAF (Wilcoxon rank-sum test, p < 0.05 for each contrast after correction for multiple comparisons, p = 0.5 for SRAF vs VPAF; Fig. 5.4H).

Figure 5.5: Tonotopic map stability over time.

(A) Registered BF maps with field boundaries from the same mouse imaged four times over a 37d period. (B) Absolute value of the BF difference computed for the same tone-responsive pixels in any pair of images. Top: Mean
Figure 5.5 (Continued)

± SEM of the BF difference from the actual maps (colored) or shuffled maps (gray). Horizontal lines indicate statistically significant differences with a Wilcoxon Sign Rank test. Bottom: Mean ± SEM of BF difference as a function of the inter-imaging interval. Individual data points are represented as circles.

Long-term stability of tonotopic maps

The chronic cranial window preparation provided long-term optical access to the ACtx, allowing us to measure the stability of frequency tuning over time. Imaging data from an example mouse conveys that macroscopic features are fairly stable over a 37-day period, while the BF of individual pixels suggests some modest session-to-session variability (Fig. 5.5A). We formalized this by registering the widefield images collected from any individual mouse and then computing the absolute value of the BF difference for all pixels that maintained frequency tuning for any given pair of imaging sessions. On average, we found that the BF for any given pixel varied by approximately 0.4 octaves between imaging sessions, which could reflect the true variability in underlying frequency tuning as well as measurement error due to image registration, threshold estimation, internal state variation, photobleaching, window clarity and myriad other experimental factors (Fig. 5.5B, top). To calculate what the BF difference would be by chance, we shuffled the pixels within each field and repeated the measurement. The average BF difference in the shuffled control decreased for smaller fields with reduced ranges of BFs, but was significantly greater than the actual BF difference in all fields (paired t-tests, t > -5.0, p < 0.000005 for each comparison). BF variability did not differ between fields (ANOVA, F = 0.37, p = 0.83) and did not systematically change as a function of interval between imaging session for any field (linear relationship between session and BF difference, p > 0.14 for all fields; Fig. 5.5B, bottom). These findings suggest that the frequency selectivity within a local patch of ACtx
(1 pixel equals ~13 µm) varies by less than a half octave over time, confirming prior reports of relative stability of adult cortical sensory maps in the absence of perturbations of sensory experience or afferent activity levels (Jenkins et al. 1990; Masino and Frostig 1996; Polley et al. 2004).

**Multiscale imaging of auditory cortex**

Whereas the mesoscale tonotopic organization of core fields in mouse ACtx has been confirmed time and again with microelectrode mapping, intrinsic signal imaging or widefield calcium imaging, the underlying organization at a cellular scale remains a point of dispute. The seminal ACtx 2-photon calcium imaging studies relied on bulk-loaded calcium dyes in anesthetized mice and reported that BFs between neighboring L2/3 cells often varied by an octave or more and only weakly conformed to a global tonotopic gradient (Bandyopadhyay et al. 2010; Rothschild et al. 2010). A subsequent study using genetically encoded calcium sensors in awake mice reported a striking precision of local frequency tuning, where the BFs of individual neurons were virtually perfectly aligned to the global tonotopic gradient (Issa et al. 2014). Recent GCaMP6 imaging studies in the ACtx of awake mice suggest that the BFs of neighboring neurons are homogeneous than what was reported in the seminal studies, although some local scatter can be qualitatively appreciated from their example images (Kato et al. 2016; Kuchibhotla et al. 2017).

We reasoned that two factors could affect the correspondence between cellular and mesoscale measurements of ACtx frequency tuning: First, we noted that the study reporting homogeneous local BF tuning did not analytically remove the influence of neuropil from the fluorescence signals measured around individual L2/3 soma (Issa et al. 2014). We expected that fluorescence arising from the surrounding axons and dendrites would reflect the aggregate
frequency tuning of the local cellular neighborhood, would more closely match the bulk widefield Ca2+ signal, and should produce more homogenous local BFs. Second, unit recordings and imaging studies have observed that some L2/3 neurons are driven by tones but have poorly defined, irregular FRAs that cannot be accurately described with a singular BF. Our prior study applied the d-prime statistic (d’) to neural FRAs and concluded that the tonotopic organization in L2/3 was substantially degraded when neurons with low d’ were included (Guo et al. 2012).

We expected that including neuropil fluorescence and selecting cells with well-defined FRAs would favor homogenous local BFs that closely matched global tonotopic BF gradients. Conversely, removing neuropil contamination and computing a BF for all tone-responsive neurons independent of tuning quality would produce more heterogeneous local BFs with a coarse global tonotopic organization. To test these predictions, we performed 2-photon imaging of L2/3 pyramidal neurons from a cohort of mice that had undergone widefield imaging 3-14 days prior (N=4, Fig. 5.6A). We spatially registered the fields of view from the tandem lens widefield microscope to the 2-photon microscope so that the tuning of individual neurons could be directly matched to the surrounding mesoscale tonotopic gradient (Fig. 5.6B). With the imaging fields aligned, we quantified frequency tuning for individual neurons before and after calcium signals from the surrounding neuropil were analytically removed from individual neural somata. Somatic FRAs with and without neuropil correction could then be compared directly to signals derived from the corresponding set of pixels from the widefield map. While some L2/3 cells showed robust frequency tuning (Fig. 5.6C), many other cells showed patchy, discontinuous frequency response areas (Fig. 5.6D). The strength of frequency tuning could be quantified by the d-prime statistic (d’), which was used to compute the statistical separability between responses near the BF versus frequency-intensity combinations far away from the BF.
Clearly defined frequency tuning (d’ > 1) was only observed in approximately half of L2/3 cells that had significant tone-evoked responses and this fraction was significantly lower in VPAF (n = 587 neurons) compared to A1 (n = 1,482), AAF (n = 2,163) and SRAF (n = 1416; Wilcoxon rank sum tests, corrected with Holm-Bonferroni, p < 1 x 10^-6 for each comparison; Fig. 5.6E). As predicted, cellular BFs in A1, AAF and SRAF more closely match the global tonotopic maps when comparisons are limited to cells with clearly defined frequency tuning (d’ > 1) and when the local neuropil contribution is not removed (Fig. 5.6F). To formalize these observations, we compared the absolute difference in BF between individual neurons to the corresponding pixels of the widefield BF map. We first defined the lower and upper bounds of the BF difference range from the chronic widefield imaging dataset reported above (Fig. 5.5). We reasoned that the difference between an individual neuron and the widefield map recorded in two different imaging sessions would not likely be smaller than an individual widefield pixel compared to itself across two imaging sessions (WF limit, solid gray line Fig. 5.6G). Conversely, the BF difference occurring by chance can be defined by the widefield BF difference between a given reference pixel and a randomly selected comparison pixel from a second imaging session (WF chance, dashed gray line Fig. 5.6G).

We analyzed frequency tuning from tone-responsive L2/3 neurons in four mice using the same criteria and analysis methods as the widefield imaging data. We confirmed that BF tuning from individual neurons was a significantly better match to the widefield map before the neuropil contribution is removed (Kruskal-Wallis for A1, AAF and VPAF, p < 0.01; SRAF, p = 0.054). After neuropil correction, the BF difference between individual neurons and the widefield signal was significantly reduced for neurons with stronger overall frequency selectivity (Kruskal-Wallis, p < 0.000001 for all fields; Fig 5.6F). For neurons with poor frequency selectivity, the
alignment to the widefield BF gradient is close to chance, consistent with an underlying heterogenous salt and pepper organization (Bandyopadhyay et al. 2010; Rothschild et al. 2010). However, for neurons with well-defined pure tone receptive fields, the alignment to the widefield map is significantly stronger and approaches the measurement limit (approximately 0.4 octaves), in agreement with 2-photon imaging of genetically encoded calcium indicators in awake mice (Issa et al. 2014, 2016). By accounting for the influence of neuropil and frequency tuning strength, our data suggest that the two ostensibly contradictory descriptions of mouse ACtx organization – locally heterogenous or locally ordered – might, to a degree, both be correct.

**Local and global organization of tonotopy at a cellular scale**

As a final step, we explicitly analyzed local BF heterogeneity from our 2-photon imaging data that had been corrected for neuropil contamination and focused on the observation that the distribution of local BFs could be substantially greater when all tone-responsive neurons are included (Fig. 5.7A, top) than when only well-tuned neurons are considered (Fig. 5.7A, bottom). To quantify local BF scatter, we applied a d’ threshold to each field of view to include all tone-responsive neurons (d’ ≥ 0) or to be restricted to neurons with increasingly high d’ values. For each d’ threshold, we treated each cell as a reference and identified all other cells within a 50 µm radial distance. We computed the median BF across all cells within a given local neighborhood and determined the BF difference for each cell relative to the median BF value (Tischbirek et al. 2019). We iterated this process for each neuron in the field of view and compiled a histogram of BF differences, using the interquartile range of the BF difference distribution as an index of local BF heterogeneity (Fig. 5.7C). Local BF scatter was significantly dependent on the strength of frequency tuning in A1, AAF and SRAF, but not in VPAF (permutation tests, p < 0.005 for A1, AAF and SRAF, p = 0.64 for VPAF, respectively; Fig.
This analysis confirms that in the core fields, estimates of local BF scatter could vary by a factor of two based solely on the inclusion criteria for selecting candidate neurons for analysis. In VPAF, by contrast, the degree of scatter is independent of frequency tuning quality and appears inherently more variable than other cortical fields.

As a corollary to local BF heterogeneity, we quantified the linear relationship between the BF of individual cells and their position along the low-to-high frequency extent of the corresponding field. Reports emphasizing local BF heterogeneity have described a very coarse correlation to global tonotopic position (Pearson R ~ 0.2), while reports of homogeneous local BF s describe a precise linear correlation (Pearson R ~ 0.9) (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Issa et al. 2014). To address these observations in our data, we projected the position of each individual cell onto a vector that connected the low and high end of the global tonotopic gradient for the corresponding field (Fig. 5.7E-F). We then computed the Pearson correlation coefficient between BF and tonotopic position. As suggested from our prior analyses, the global correlation strength, operationally defined here as the bootstrapped Pearson correlation coefficient, could be accurately described as coarse or precise, depending on the inclusion criteria for allowing tone-responsive neurons into the analysis. Global correlation strength was significantly greater for neurons with higher d’ values in A1, AAF and SRAF, though again not in VPAF, where cellular tuning did not conform to an overall tonotopic scheme (Permutation test, p < 0.001 and p = 0.77, respectively; Fig. 5.7G).
Figure 5.6: Spatial alignment of mesoscale and cellular frequency tuning.

(A) Schematic of widefield and 2-photon imaging systems. (B) Top: Field of view (FOV) registration between the widefield and 2-photon imaging systems. Bottom: BF widefield map from an example mouse with 2-photon imaging FOVs superimposed. (C) A cell (orange outline) identified with an arrow in B is shown with 2-photon excitation. The frequency response area are derived from the z-scores of 2-p GCaMP6s signals after neuropil correction (soma) before neuropil correction (S+N) and from the matching pixels in the widefield system (WF). Mean ± SEM of the fractional change in fluorescence identified by the analysis software arising from the soma or from neuropil. (D) An additional cell recorded in with weaker, patchy tuning for frequency and lower d’ values. Blue outlines in 6D-E identify the five reference frequency-level points used for the calculation of d’. Downward arrows indicate the BF of the corresponding pixels from the widefield image. (E) Cumulative density histograms of d’ values in each cortical field. (F) Widefield BF maps are extracted from rectangular regions of interest identified in three different mice. Widefield tonotopy is presented in the background with superimposed 2-photon BF's from individual cells. Cellular
Figure 5.6 (Continued)

BF is shown before (left) versus after (right) neuropil correction and for all sound-responsive neurons (top) versus only neurons with FRA d’ values greater than 1.0 (bottom). (G) Absolute value of the BF difference between individual neurons and the underlying widefield map are shown for four cortical areas as a function of frequency tuning strength. Dashed gray line represents the mean ± SEM of the BF difference that would occur by chance. Solid gray line represents the mean ± SEM of the smallest possible BF difference, defined from the difference between widefield vs widefield imaging sessions shown in Figure 5.

Discussion

In this study, we performed multiscale imaging from a transgenic mouse that expressed a genetically encoded calcium indicator in excitatory neurons throughout the cerebral cortex but retained excellent hearing into adulthood (Fig. 5.1). We described a procedure to pre-process the widefield epifluorescence signals and identify an auditory response threshold and BF for each pixel individually (Fig. 5.2). We marked the caudal, rostral, medial and lateral extremes of the ACtx with Di-I and noted that the lateral edge of the tonotopically organized map extended farther caudally and laterally than would be expected from the anatomical landmarks identified in widely used mouse brain atlases. We implemented a data-driven parcellation approach that used abrupt drops in signal strength and reversals or phase shifts in BF gradients to identify five cortical fields – A1, AAF, SRAF, VPAF and DP (Fig. 5.3). We observed statistically significant mesoscale tonotopic organization in all fields, where modules of similar frequency tuning bandwidths and response thresholds were superimposed on BF gradients (Fig. 5.4). Frequency tuning remained relatively stable over approximately one month of repeated imaging from the same mice, where the BF of individual pixels varied by less than 0.5 octaves (Fig. 5.5). We used 2-photon imaging to spatially register the frequency tuning of individual neurons to the widefield maps. We found that correspondence between the cellular and mesoscale tonotopic maps ranged from slightly better than chance to nearly equal to the measurement limit, depending on the
strength of cellular frequency tuning and whether neuropil contributions were factored out or included (Fig. 5.6). With the neuropil contamination removed, we found that the degree of local BF scatter and orderly progression of local BFs along the tonotopic axis both reflected the strength of frequency tuning (Fig. 5.7).

Figure 5.7: Local and global organization of tonotopy.

(A-B) 2-photon FOV showing cellular BFs from all sound-responsive neurons versus neurons with strong frequency tuning (d’ > 1, B). For a given reference cell, all neighboring cells within a 50 μm radius are identified. The median BF for all cells within this local neighborhood is computed. (C) The difference in BF for each cell versus the neighborhood median is shown for all sound-responsive neurons (black) or for neurons that are strongly tuned to frequency (d’ > 1, red). The interquartile range of the BF difference histogram (shaded regions) provides an estimate of local BF heterogeneity. (D) Mean ± SEM BF heterogeneity as a function of frequency tuning strength for four cortical areas. (E) Widefield BF map from an example mouse with superimposed 2-photon cellular tuning
Figure 5.7 (Continued)

from two FOVs in AAF. Gray line indicated global tonotopic vector. L = low. H = high. (F) Individual cells across AAF FOVs in all mice are assigned to a position in the global tonotopic vector. Median (horizontal line) and interquartile range of BFs within each AAF position bin. The linear relationship between the mesoscale tonotopic position and cellular BF is indexed by the Pearson’s R value. (G) Mean ± SEM of the bootstrapped Pearson’s R as a function of frequency tuning strength for four cortical areas.

**Precision of frequency tuning in the mouse auditory cortex**

Tonotopy is among the most rudimentary aspects of ACtx organization and is clearly predictive of frequency guided auditory behaviors (Jenkins and Merzenich 1984; Znamenskiy and Zador 2013). Yet, the ostensibly straightforward question of whether the ACtx is tonotopically organized has remained a point of contention for over fifty years (for review see Kanold et al. 2014). Opposing papers have debated whether the tonotopic gradient in core fields was coarse or precise from the start of the modern neurophysiological era (Evans and Whitfield 1964; Merzenich and Brugge 1973) and continuing right up to the current era (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Issa et al. 2014). A consensus view is emerging in the rodent somatosensory and visual cortex for a globally ordered topographic mapping of the receptor epithelium comprised of locally heterogeneous tuning at a cellular scale (Ohki et al. 2005; Sato et al. 2007; Bonin et al. 2011; Clancy et al. 2015). In ACtx, there is clear support for a robust tonotopic organization in the core auditory fields when maps are made at mesoscale resolution, where individual points of measurement reflect a pooling of local activity. Whether studied with microelectrode multiunit recordings from the middle layers of anesthetized animals (Hackett et al. 2011b), widefield imaging of intrinsic signals (Moczulska et al. 2013; Kato et al. 2016; Tischbirek et al. 2019) or genetically encoded calcium sensors from the upper
layers of awake animals (here and in (Issa et al. 2014; Liu et al. 2019)), a highly ordered systematic progression of preferred frequency is readily apparent in the core fields of the mouse ACtx, in keeping with descriptions in over 20 other mammalian species (Kaas 2011).

Zooming into any local cellular neighborhood within the tonotopic map reveals a considerable heterogeneity of preferred frequency between neighboring single neurons, as revealed by single unit spiking of neighboring units (South and Weinberger 1995) or suprathreshold calcium events recorded with 2-photon imaging (Bandyopadhyay et al. 2010; Rothschild et al. 2010; Winkowski and Kanold 2013; Liu et al. 2019). The balance of globally ordered, smoothly changing gradients built from heterogeneous, inherently ‘noisy’ local processing is consistent with reports of globally ordered topographic mapping of MGB thalamocortical projections (Hackett et al. 2011b), yet heterogeneous frequency tuning within individual MGB axon terminals (Vasquez-Lopez et al. 2017) or between individual spines of L2/3 pyramidal neurons (Chen, Leischner, et al. 2012).

The only remaining point of debate is centered on the degree of local heterogeneity in preferred frequency tuning. On one end of the spectrum, 2-photon imaging from layer 2/3 neurons in awake mice that express GCaMP3 suggested extremely low variance in preferred frequency between local neurons in A1 (standard deviation of 0.4 octaves for a 230x230 µm field of view) and a tight correspondence between local preferred frequency and the global tonotopic gradient (correlation coefficient > 0.85) (Issa et al. 2014). However, this report did not analytically correct for neuropil contribution. Even in visually isolated soma, calcium signals recorded with 2-photon microscopes average fluorescence across tens of microns in the z-axis, where much of this signal will reflect a contribution from out-of-focus axons and dendrites from other neurons. As we show here, including the neuropil substantially reduces local BF
heterogeneity and improves the alignment to the global tonotopic map (Fig. 5.6). Therefore, we would argue that this study may have over-estimated the degree of local BF precision by including a more global neuropil signal (Lee et al. 2017). On the other end of the spectrum, the original pioneering 2-photon imaging studies of mouse ACtx described substantial heterogeneity from neuropil-corrected L2/3 cells in anesthetized animals. Increased heterogeneity is unlikely to arise solely from the anesthetized state, as systematic differences in BF tuning are not observed in cells recorded in both anesthetized and awake conditions (Guo et al. 2012; Tischbirek et al. 2019). Some of the heterogeneity reported in these studies could reflect the use of bulk-loaded dyes, which would increase the contribution of fluorescence signals from non-neuronal cell types and certain upper layer GABAergic interneurons that may have broader frequency selectivity than pyramidal neurons (Li et al. 2014). Another possibility is that these seminal studies either used tone bursts with very short inter-trial intervals (0.4s, (Rothschild et al. 2010)) or longer (1s duration) amplitude-modulated tones (Bandyopadhyay et al. 2010). Recording from neurons in an adapted state or from neurons responding during sustained periods of stimulus-related activity could increase the complexity of auditory tuning and increase the heterogeneity between neighboring neurons (South and Weinberger 1995; Wang et al. 2005; Kato et al. 2017) or may conflate the independent frequency tuning for sound onset versus offset (Liu et al. 2019).
Figure 5.8: Mapping the core and higher-order fields of the mouse auditory cortex.

(A) Schematic of the original proposal for the layout of the mouse ACtx fields (Stiebler et al. 1997). Thick black line represents the areal outline of the complete ACtx studied here. (B) BF maps from six individual mice. (C) Left: Cartoon of the general BF gradients and field boundaries suggested from the GCaMP6 imaging data reported here. Right: low-to-high frequency gradients (L and H, respectively) and field naming designations proposed here. (D) Low-to-high frequency gradients and field naming designations proposed from earlier GCaMP3 widefield imaging studies (Issa et al. 2014, 2016). Striped region denotes a tone-insensitive region identified by these studies. (E) Voronoi tessellation of preferred frequency and field naming designations from a high-density multiunit microelectrode mapping study (Guo et al. 2012). (F) Preferred frequency and field naming designation identified from a lower density microelectrode mapping study (Joachimsthaler et al. 2014). (G) Cortical parcellation scheme suggested from flavoprotein widefield imaging studies (Tsukano et al. 2015, 2016; Tsukano, Horie, Takahashi, et al.)
Figure 5.8 (Continued)

2017). Note that all schematics have been adapted, re-annotated and resized from their original form to fit the right hemisphere.

The degree of local heterogeneity reported here is somewhere in between these two extremes and is qualitatively consistent with our prior descriptions from electrophysiological recordings of single units (Guo et al. 2012) and a recent demonstration of cellular tuning across all layers of the A1 column (Tischbirek et al. 2019). The most important point from the data presented here was that correspondence to the widefield map (Fig. 5.6G), the degree of local BF heterogeneity (Fig. 5.7D) and the global correlation (Fig. 5.7G) could differ by a factor of two, depending on the inclusion of neurons that were activated by tones, but were poorly selective for a single, narrow range of frequencies. This is completely self-evident and corresponds exactly to the same observation we made previously in electrophysiological recordings from L2/3 units (Guo et al. 2012). Essentially, if a neuron had an irregular, broad or multi-peaked receptive field, its frequency preference cannot be as accurately reduced to a single number. Regardless, even by limiting our sample of neurons to those with reasonably strong frequency selectivity, it was clear that the BFs of neighboring neurons do not reflect a salt and pepper organization, but rather are strongly predicted by their position within the overlying mesoscale map, but vary on the order of approximately half an octave in any local neighborhood (~50 µm XY radial distance). This degree of local heterogeneity could be an unavoidable consequence in small brains with substantial divergence of thalamocortical and intracortical connectivity (Hackett et al. 2011b). Reduced heterogeneity would be expected along a radial column than along the tangential plane, or in species with larger auditory cortices or more precise anatomical connectivity (Atencio and Schreiner 2010, 2013; Guo et al. 2012; See et al. 2018; Tischbirek et al. 2019), but we expect to
see that globally systematic yet locally heterogenous selectivity would be an organizing feature of sensory cortex organization (Kanold et al. 2014).

**Organization and naming schemes for multiple fields of the mouse auditory cortex**

Whereas there is general agreement about the balance between global order and local diversity in A1 and AAF, there is no consensus on what fields of the mouse ACtx are “higher-order”, where they are located or even what they should be called. Here, we implanted a cranial window to cover the full extent of the mouse ACtx, affording us optical access to lateral areas of the cortex that are difficult to record from with acutely inserted microelectrodes (Fig. 5.8A). Looking across individual mice (Fig. 5.8B), we consistently observe four low-frequency hubs at the edges of the ACtx that fan out and collide with one another to form the boundaries between fields (Fig. 5.8C, left). In keeping with the seminal mapping study as well as the nomenclature adopted in other species, the objective parcellation approach used here is consistent with having A1, AAF and DP labeled as separate fields (Fig. 5.8C, right). The frequency gradients identified by our mesoscale GCaMP6 imaging are a close fit to that described by earlier mapping of GCaMP3 signals with one exception: They argue that there is a tone-insensitive region at the border of A1 and AAF (Fig. 5.8D, (Issa et al. 2014, 2016)). We confirmed that pixels in this area have higher response thresholds, but with the individual pixel thresholding procedure used here it is evident that even though tone-driven responses are weaker in this region, they have BF that are consistent with the overall frequency gradients linking A1 and AAF, in keeping with over 20 other mammalian species and with prior microelectrode mapping of the A1-AAF junction in the mouse (Linden et al. 2003; Hackett et al. 2011; Kaas 2011; Guo et al. 2012; Shepard et al. 2015). Prior calcium imaging (Fig. 5.8D) and microelectrode mapping studies also identified an area with low-frequency BFs lateral to A1 and AAF (Fig. 5.8E and 5.8F, respectively (Stiebler et al.
By developing a surgical approach to position the cranial window more laterally and caudally than previous studies, we identified a fourth low-frequency hub that we named VPAF, in keeping with a description of a similar ventral-posterior field identified with widefield imaging of intrinsic signals in the rat ACtx (Kalatsky et al. 2005).

Although naming conventions vary, the overall tonotopic gestalt is consistent across these widefield calcium imaging and microelectrode mapping studies. The only exception comes from a series of reports using widefield imaging of endogenous flavoprotein signals that describe a low-frequency area interposed between A1 and AAF and mis-identify the upper limb of A1 as belonging to a separate field referred to as DM (Tsukano et al. 2015, 2016; Tsukano, Horie, Ohga, et al. 2017; Tsukano, Horie, Takahashi, et al. 2017) (Fig. 5.8G). Flavoprotein fluorescence signals are an order of magnitude slower and weaker than genetically encoded calcium sensors. Possibly on account of the need for longer trial durations, these studies generally test a more limited set of tone frequencies at a single sound level, as compared to the 72 frequency/level combinations used here. As tonotopy is substantially degraded when derived from tones presented at a single suprathreshold sound level (Fig. 5.4C and (Guo et al. 2012), the organizational scheme suggested from these studies should be interpreted cautiously.

Mouse ACtx researchers have generally adopted the original naming scheme proposed by the seminal low-density microelectrode mapping studies of the mouse ACtx (Fig. 5.8A and 5.8F). We propose that some aspects of the original scheme are misleading and should be changed in favor of a naming system that is more consistent with auditory cortical fields in other mammals. The designation of an ultrasonic field (UF) should be abolished on the grounds that i) “ultrasonic” is anthropomorphic and refers to any frequency above the limit of human hearing
(approximately 20 kHz) and could therefore refer to approximately half of the mouse hearing range and not just the representation of frequencies above 45 kHz; ii) there doesn’t appear to be anything discrete or discontinuous about the cortical representation of frequencies above 45 kHz; in most cases it simply appears as the high frequency elaboration of the A1 and AAF tonotopic gradients, in keeping with the elaboration of high-frequency BF's described in the tonotopic representations of the mouse central nucleus of the inferior colliculus (Garcia-Lazaro et al. 2015).

Further, we argue that A2 is a misnomer for the field described here as the suprarhinal auditory field. There was never a particularly compelling reason to label this area as A2 to begin with. The seminal study relied on low-density microelectrode mapping to note the frequency tuning bandwidths were wider and the tonotopic organization less clear than in the core fields (Stiebler et al. 1997), both of which have been confirmed here (Fig. 5.4). The designation of a secondary auditory field is only widely used in cats but is most comparable to a non-tonotopically organized parabelt areas in non-human primates (Reale and Imig 1980; Schreiner and Cynader 1984). In other rodents, carnivores and non-human primates, field naming conventions follow the anatomical position of the field and not its presumed position in a hierarchy of cortical processing. Among the commonly used animal models for ACtx research, laboratory rats (Rattus) are the closest evolutionary relative to the laboratory mouse (Mus). Because the suprarhinal auditory field described here shares the same position and tonotopic orientation as the rat, we argue that this field should also be called SRAF (Polley et al. 2007). Perhaps more importantly, higher-order auditory fields – by definition – receive their predominant thalamic input from higher-order thalamic subdivisions (e.g., MGBd), not primary thalamic subdivisions (MGBv) (Rose and Woolsey 1949a, 1949b; Andersen et al. 1980; Winer et
Dual neuranatomical tracer injections into A1 and SRAF in the mouse revealed that both fields receive inputs from separate zones of the MGBv, with hardly any input from MGBd (Ohga et al. 2018). Although a second study that did not reconstruct the full rostral-caudal extent of the MGB and did not use physiological guidance for their tracer injections came to the opposite conclusion (Ji et al. 2016), at a minimum the source of thalamic input to these areas is uncertain and deserves further study with physiologically guided tracer injections before any attribution of a primary or secondary level of processing can be made.

Where – if anywhere – are the higher order fields of mouse auditory cortex?

Unlike the visual cortex, where hierarchies for stimulus processing abound, differences in the nature and form of auditory stimulus processing between fields of the ACtx are more a difference of degree than a difference of kind. Three notable exceptions have been identified. First, strictly non-primary areas of the ACtx have been identified in the human brain, where lateral regions show specialized responses for music and speech that are not observed in the primary areas (Leaver and Rauschecker 2010; Norman-Haignere et al. 2015; Overath et al. 2015; Kell et al. 2018; Norman-Haignere and McDermott 2018). Second, higher-order areas have been identified in the ferret ACtx that selectively encode sounds according to their behavioral meaning and not their acoustic features (Atiani et al. 2014; Elgueda et al. 2019). Third, a sub-type of neuron with broad spike waveforms have been identified in a higher-order field of songbird ACtx that supports the de novo emergence of sparse, contrast-invariant representations of conspecific vocalizations (Schneider and Woolley 2013; Kozlov and Gentner 2016; Ono et al. 2016).

Here, we used pure tone bursts in passively listening mice to delineate the boundaries of cortical fields without revealing much about any underlying specializations. Mesoscale tonotopy
was strong in A1, AAF and SRAF (Fig. 5.4C), where underlying neurons that were well-tuned to sound frequency showed comparably homogeneous BFs (Fig. 5.7D) and adherence to global tonotopic vectors (Fig. 5.7G). VPAF and DP exhibited less organized, incomplete representations of the cochlear frequency map. Cellular imaging in VPAF revealed a highly disorganized salt and pepper organization of BFs that was largely insensitive to frequency tuning strength.

Ultimately, differences in sound frequency organization can suggest candidates for higher-order fields, but cannot provide definitive evidence for where a field sits within a cortical hierarchy or heterarchy. “Higher-order” is an anatomical designation that reflects a preponderance of higher-order thalamic inputs, stronger inputs from brain areas that encode multi-sensory inputs, and stronger connectivity with non-auditory structures such as frontal cortex, amygdala or neuromodulatory centers. Functional markers such as tonotopic precision, task-related modulation and cross-modal sensitivity are products of – rather than determinants of – the complement of afferent inputs that define a region as core or higher-order. In this regard, the poorly selective disorganized frequency representations in VPAF and DP suggest a relatively weak input from the MGBv and stronger input from higher-order brain areas, but this can only be demonstrated with carefully positioned injections of tracers or viral vectors. In mouse visual cortex, researchers have made good headway identifying specialized visual feature processing and state-dependent modulation in fields beyond V1 (Glickfeld et al. 2014; Ramesh et al. 2018; Beltramo and Scanziani 2019). In mouse ACtx, researchers have by and large focused on recording only from A1, typically in passively listening animals. Here, we propose that VPAF and DP may represent good candidates for studies on higher-order anatomical connectivity.
studies as well as neurophysiological experiments that focus on the extraction of auditory features that guide purposeful behavior.
CHAPTER 6

CORTICAL MAP DYNAMICS ASSOCIATED WITH PERCEPTUAL HYPERSENSITIVITY FOLLOWING ACOUSTIC TRAUMA

Ariel Edward Hight, Matthew McGill, Kameron K. Clayton, Aravind Parthasarathy, Kenneth E Hancock, & Daniel B Polley

ABSTRACT

Cortical neurons remap their sensitivity to topographically adjacent areas following restricted peripheral lesions. To address how this reorganization at the cortex relates to changes in perception, we leveraged chronic calcium imaging in awake mice to track mesoscale map reorganization in excitatory neurons across fields A1, AAF, and SRAF of the auditory cortex following a novel sudden high-frequency acoustic over-exposure in mice. Peripheral measures of cochlear integrity and function reveal high-frequency permanent threshold shifts (PTS) in ABR wave 1b (40-60 dB SPL) primarily due to loss of outer hair cell amplification. Operant tone detection confirmed hearing loss in the region of PTS. Although peripheral measures revealed no damage or changes in low-frequency coding, we found an enhanced behavioral detection of tones at these spared frequencies. We found that day-by-day measures of cortical reorganization and plasticity were tightly related to behavioral enhancement. Additionally, the magnitude of recovery revealed distinct time-courses for plasticity across the different cortical area. Taken together, these findings suggest perceptual consequences following sensory trauma reflect central reorganization and differences in cortical reorganization between cortical areas likely
reflect unique microcircuits, short range horizontal circuits, long range projections from higher-order areas, or both.

INTRODUCTION
The brain employs a remarkable set of controls for both homeostatic and hebbian plasticity. The balance between the two maintain neural activity around a set point while permitting new excitatory connections to maintain circuit dynamics such as learning (for review, see Keck et al. 2018). Sudden peripheral trauma engages these control systems at their limits as large regions of neurons suddenly lose their primary input. The first steps towards homeostatic balance following peripheral damage are associated with rapid disinhibition (within hours) as demonstrated in visual, somatosensory, and auditory systems (Calford and Tweedale 1988, Gilbert and Wiesel 1992, Goel and Lee 2007, Chen et al 2012 Chen et al 2011 Keck et al 2011, Kuhlman et al 2013 Li et al 2014, van Versendaal et al 2012, Sammons and Keck 2015, Resnick and Polley 2017). Other mechanisms for homeostatic balance include synaptic scaling (Turrigiano et al. 1998), constraints on spine growth (Matsuzaki et al. 2004; O’Donnel et al. 2011), and spine fluctuations (Yasumatsu et al. 2008). How these mechanisms work in concert with one another and how disparities in their time and spatial scales may be reconciled are important for understanding how these self-regulating processes operate and how they may shape sensory processing dynamics. For example, sensory coding relies on a proper excitatory-inhibitory balance (Isaacson and Scanziani 2011; Shadlen and Newsome 1998; Pouile et al. 2009; Zhou et al. 2014) and maintenance of network stability is effective but not overbearing as enable plasticity (Vogels et al. 2011; Froemke et al. 2013; Hellyer et al. 2016; Barret et al. 2016). Wholesale cortical
reorganization following hearing loss will require an orchestrated array of homeostatic elements across the functional organization of the auditory cortex.

The organization of frequency in the auditory cortex is tightly coupled with the perception of sound. Cortical lesions restricted to a set of frequencies lead to deficits in learned behaviors requiring the same frequencies corresponding to but not away from the lesioned frequencies (Jenkins and Merzenich 1984). Perceptual learning and training are correlated with persistent cortical expansion and reorganization (Recanzone et al 1993; Blake et al. 2002; Rutkowski and Weinberger 2005; Witte and Kipke 2005; Polley et al 2006). Areas of the auditory cortex are markedly different; varying degrees of attentional and behavioral modulation are found in other topographical areas than A1 (Atiani et al. 2014) and reflect unique patterns in receiving top-down projections. In line with evidence for cortical involvement in the perception of sound, recent evidence suggests some perpetual disorders arise from central plasticity following peripheral damage such as drug and noise-induced hyperacusis (Auerbach et al. 2014, Auerbach 2019), noise-induced tinnitus (Yang et al. 2011; Auerbach 2014), and temporal processing deficits in auditory neuropathy spectrum disorder (Chambers et al. 2016a).

Noise-induced hearing loss (NIHL) affects up to a quarter of adults under 70 and usually manifests as damage in high frequency sensitivity (NIH.gov). Hearing loss such as NIHL, high-frequency hearing loss (HFHL) and other types lead to difficulty in speech comprehension and other debilitating symptoms such as tinnitus and hyperacusis. Here, we introduce a model for sudden high-frequency hearing loss in adult mice that that result in common profiles of high-frequency hearing loss as found in humans. Immediately following this acoustic trauma we found an enhancement in low-frequency sound driven behavior, which is paradoxical given that our measures of integrity and function at the low-frequency region of the periphery appear
normal. Our prediction of central plasticity following hearing loss was confirmed by our day-by-day widefield epifluorescence imaging of genetically encoded calcium indicator GCaMP6 across excitatory neurons of the auditory cortex in mice. Our observations of reorganization and plasticity were consistent across cortical areas A1, AAF, and SRAF. Moreover, time-course of reorganization and plasticity paralleled enhancements in tone detection behavior of the same frequency. These results suggest that enhancement of tone detection reflects plasticity in the central auditory pathway and that differences in reorganization between areas of the auditory cortex reflect differences in short range horizontal circuits, long range projections from higher brain areas, or a combination of both.

METHODS

Experimental model and subject details.

Imaging and Behavior experiments were performed on adult offspring of Thy1-GCaMP6s mice (Jackson labs stock number 025776) crossed with CBA/CaJ mice, which retain good hearing into adulthood (Zheng et al. 1999, Romero et al. 2019). Offspring of both sexes were implanted with a head plate for head-fixation. All offspring were affixed with head-fixation hardware. Offspring positive for the thy1-GCaMP6 genotype were used for imaging and were implanted with a cranial window in addition to head-fixation hardware, during the same procedure. These procedures were performed at or after 6 weeks while allowing time to recover before undergoing acoustic trauma before 9 weeks. Mice were housed cooperatively prior to head-fixation and housed individually thereafter. Mice were provided *ad Libitum* access to food and water, unless noted otherwise (behavior), and were maintained in 12-h day/night cycles. All procedures were approved by the Massachusetts Eye and Ear Infirmary Animal Care
and Use Committee and followed the guidelines established by the National Institute of Health for the care and use of laboratory animals.

Figure 6.1: High-frequency noise exposure induces high-frequency PTS caused by loss of OHC function

(A) Narrow-band (16-32 kHz) noise exposure is presented at 103 dB SPL for 2 hours via a calibrated free-field speaker. (B) Mean +/- SEM ABR wave 1b thresholds measured across 4-65 kHz prior to and at days post-exposure 1, 7, and 10 (n = 36 ears, n = 16 mice) (C) Mean +/- DPOAE thresholds across frequency and days post exposure (same mice and days as in (B)) (D) Counts of individual synapses are identified by co-localization of presynaptic ribbons (red) and postsynaptic receptors (green) Scale bar: 10 µm (E) Counts of outer hair cells are identified across
three rows by the presence or lack of (dashed line) adjacent cells. Scale bar: 10 µm (F) Immunohistological counts of synaptic ribbons per inner hair cell (IHC) in noise (n = 3) and sham (n = 3) exposed mice are plotted as a function of position along cochlear length. (G) Mean cytocholeograms indicate outer hair cell (OHC) counts as a function of position along cochlear length, in the same ears used in (F). (H) Left: Micrographs of OHC stereocilia stained with espin, in noise (top) and sham exposed (bottom) mice. Arrowheads indicate location of stereocilia disarray in the first row of OHCs. Cochlear place is approximately 32 kHz. Scale bar: Right: zoom-in image of a single OHC from first row in noise (top) and sham (bottom) exposed mouse.

**High-frequency noise exposure**

Octave-band noise at 16-32 kHz were presented at 103 dB SPL for 2 hours. Exposure stimuli was delivered via a tweeter (Fostex USA) fixated inside a custom-made exposure chamber (20 x 20 x 20 inches). The walls of the acoustic enclosure were slanted such that no two walls were parallel as to minimize standing waves. Additionally, irregular surface depths and sharp edges were built onto 3 of the 6 walls using stackable ABS blocks to diffuse the high-frequency sound field. Calibrated measurements at the center of the chamber mimicking the position of a mouse during exposure, showed a maximum variation of 2-3 dB SPL. Prior to exposure mice were placed, unrestrained, in an independent wire-mesh chamber (6 x 6 x 4 inches). This chamber was placed at the center of a continuously rotating plate, ensuring mice were exposed to a more uniform sound field. Sham exposed mice underwent the same procedure except the speaker was turned off prior to exposure.

**Auditory brainstem response (ABR) and distortion-product otoacoustic emission (DPOAE) measurements.**

Mice were anesthetized with an initial dose of xylazine and ketamine and were subsequently boosted with ketamine to maintain a stable plane of anesthesia during the entire duration of
testing. Core body temperature was measured with a rectal thermometer and maintained at 36.5 degrees with a homeothermic blanket system. All sound pressure presentations and recordings were facilitated by a custom acoustic assembly consisting of two electrostatic drivers and a miniature microphone. For ABR measurements, tone bursts (2-64 kHz, 0.5 octave steps, 5 ms duration, 0.5 ms rise-fall time, alternating lead polarity) were delivered at 27 Hz and at 5 dB increments from 20-80 dB or up to 100 dB SPL for high-threshold responses. Recordings were made from subdermal needle electrodes positioned immediately ventral to each pinna (+ and -) and the ground electrode placed at the dorsal surface, near the tail. Filtered ABR recordings (gain = 10k, bandpass filter = 0.3-3 kHz, trial average = 1024) were thresholded based by determining the lowest intensity at which clear ABR wave 1b peaks were identified and was also present at higher intensities. DPOAEs were recorded from the ear canal, amplified, and digitally sampled. Fast Fourier transform (FFT) magnitudes were computed and then averaged by waveform and spectral averaging. Thresholds were defined as the f1 level necessary to produce a identifiable spectral peak at 2f1-f2. Both DPOAE and ABR measurements were taken for each ear and during the same recording session. For imaging animals, these measures of cochlear function were measured prior to noise exposure and post-exposure days 1, 7, and after all imaging sessions have been completed. A separate group of 5 mice were used to assess the efficacy of noise-exposure, cochlear function in this group was measured pre-exposure and at post-exposure days 1, 3, 5, and 7. Behavioral animal cochlear function was measured at the conclusion of behavior data collection.

*Detailed surgical procedures for head-fixation hardware and cranial windows.*
We used a procedure previously described in full detail (Romero et al. 2019) and is described here briefly. Mice 6 weeks old of either sex was anesthetized with isoflurane in oxygen (5% induction; 1.5–2% maintenance). After removing the periosteum from the dorsal surface of the skull, a custom titanium head-plate (iMaterialise) was bonded to the dorsal surface of the skull with dental cement (C&B Metabond). For cranial windows, previously applied dental cement is allowed to harden for 30 minutes while the tendons of the dorsal aspect of the temporalis muscle over the auditory cortex are dissected and displaced laterally, where the exposed surface was hardened and sealed with cyanoacrylate (Vetbond, 3M). After the dental cement hardens, a 3mm diameter craniotomy is carefully positioned and performed. The exposed cortex is rinsed with saline and a prepared 3mm+4mm window stack is placed into the craniotomy with the 4mm creating a seal with the bone. While maintaining positive pressure on the window, any exposed gaps were sealed with a silicone adhesive (Kwik-Sil) and then dental cement is then applied to seal the window and all exposed surfaces. After the cement hardened, post-operative injections of Buprenex (0.05 mg/kg) and Meloxicam (0.1 mg/kg) were administered before the animal was allowed to recover in a warmed chamber.

**Preparation for behavioral testing.**

Mice were allowed 1-2 days for recovery before beginning water-restriction, where water bottles were removed from their cages and they were given 1 mL of water per day. Once mice had reached 80% of their previous body weight, behavioral training was started.

**Tone detection behavioral measurements.**
Mice were trained in a go-no go tone detection task with interleaved catch trials to measure false-alarm rates. Mice were head-fixed in a custom-built, sound-insulated rig, where tones were presented via a free-field speaker positioned 10 cm from the left ear. Mice indicated their decision by licking a stainless steel lickspout (vetspecialtyproducts.com). Once a tone (100 ms) was presented, mice were given 1.5 seconds after tone onset to lick the spout for a small water reward. Licking during catch trials was punished with a 5 second timeout. Inter-trial intervals were varied between 3 and 10 seconds based on an exponential distribution. Mice were trained to detect 8 kHz and 32 kHz tones in a 2-down, 1-up adaptive tracking paradigm, where two correct detections were required to decrease the sound intensity by 5 dB SPL and one miss was required to increase the sound intensity by 5 dB SPL (figure 6.2c). This method enabled a stable day-by-day measurement of detection threshold. A modified version of this adaptive tracking method was used where, after each change in intensity level, 3 trials were presented: +/- 5 dB SPL relative to the last intensity tested, and a catch trial. Enough trials to construct a psychometric curve was enabled using this method. A single tone was presented until 1 reversal was reached, and then the other tone was presented; a run was completed once 6 reversals had been reached for both frequencies. Once consistent performance on this task was reached, indicated by steady threshold estimates across 3 or more days and false-alarm rates below 30%, 3 days of baseline behavioral performance were collected. Mice were then either noise- or sham-exposed and three weeks of behavioral testing were conducted: 0- 8, 10, 12, 14, 16, 18, and 20 days after exposure. Noise exposures were conducted on 7 mice and sham exposures on 6 mice.
Figure 6.2: High-frequency hearing loss confirmed by operant behavior via adaptive tracking with flankers paradigm.

(A) head-fixed and mice are trained in a go/no-go task. High (32 kHz) and low (8 kHz) frequency tones are presented in free-field and mice report sound with a lick. Correct responses are rewarded with water. Timeline indicates days of operant behavior data collection relative to the day of noise (n=7 mice) or sham exposure (n=6 mice). (C) Adaptive tracking determines the sequence of tone presentation intensity. Consecutive correct responses result in lowering of tone intensity, incorrect responses result in increasing tone intensity. This two-down, one-up
Figure 6.2 (Continued)

procedure (see solid line) is flanked by tone presentations at +/− 5 dB SPL (see individual trials flanking solid line). Catch trials are shown in grey. (D) Raw hit rates and corresponding logistical fits are plotted on representative days (rows) for low and high frequency tones (columns). (E) Mean ± SEM of detection thresholds in noise and sham exposed mice as a function of day re. exposure. (F) Scatterplot of thresholds for behavioral detection (y-axis) and ABR wave 1b (x-axis) are plotted for noise and sham exposed mice, pre and post-exposure.

Data analysis of behavioral responses.

To estimate threshold for an individual mouse, all reversals across runs for a single day were averaged for each frequency. To obtain functions for describing detection performance across intensity for each frequency, raw (hit or miss) data was first concatenated across a group of days excluding runs with false-alarm rates above 40%. To properly concatenate the data across days, intensities were converted to dB re. threshold, meaning relative to the threshold estimate for that day. In this grouped data, intensities with less than 5 trials were excluded from the analysis. On this grouped data, specifically probability of hit versus dB re:threshold, a logistic regression was performed:

\[
\text{logit} (p(\text{hit})) = \beta_0 + \beta_1 \times \text{intensity}
\]

From this regression, \( \beta_1 \) was used as the slope measurement for the behavioral detection function. Slope measurements were normalized individually for each mouse relative to their baseline performance.

In-Vivo calcium imaging.
Imaging microscope, stimulus presentation, and data acquisition system employed here is the same as described in Romero et al. 2019, except for the number of trials presented and days of imaging. Tone bursts (50 ms, 5ms on-off ramp) were presented at 4-64 kHz, 0.5 octave steps and sound intensity varied from 0-70 dB SPL in 10 dB steps. Trials were 3.5 seconds in duration and each tone-intensity combination was repeated 25 times. **Timeline of Imaging:** 8 noise-exposed and 6 sham-exposed mice underwent three days of pre-exposure imaging. After exposure, mice were imaged for consecutive days (days 0-5), every other day (days 7-30), and then every week (days >30) until the window quality was degraded.

**Image processing**

The step-by-step reduction of baseline fluorescence to dF/Fo is the same as described in Romero et al. 2019 and is described briefly here. First, the slowly varying baseline of the raw fluorescence was computed using a detrending algorithm (Chronux toolbox, http://chronux.org/). Next, dF/Fo was reduced from the raw trace by and dividing the difference between the raw fluorescence and baseline by the baseline, (raw fluorescence - baseline)/baseline. The trial averaged dF/Fo trace was further reduced to a single magnitude by subtracting the baseline dF/Fo from the weighted average between the peak magnitude and the magnitudes before and after the peak. The dF/Fo baseline was determined by the mean of the trial averaged trace prior to tone onset, t=0-0.45 s and the peak of the time averaged traced was determined by the maximum in the 0.5 s time window following tone onset. A 4x4 median filter was applied to the resulting spatial distribution of magnitudes evoked by tones at each frequency and intensity combination.
Reduction of magnitude to cortical maps:

Pixel-by-pixel thresholds to sound were computed day-by-day and level-by-level. The distribution of evoked magnitudes under the cranial window were computed across frequencies individually for each level and all pixels with a magnitude below 65% of the cumulative distribution were eliminated. Next, the remaining pixels were eliminated if they were not a) spatially adjacent to an active pixel, b) at an adjacent frequency for the same pixel location, and c) active at an adjacent intensity at the same pixel location. Pixel thresholds were determined by the lowest intensity in which all previous threshold criteria were met. Preferred frequency for each pixel was determined by first taking the mean at each frequency, across intensities from respective pixel threshold to 20 dB above threshold. The resulting distribution of mean magnitudes across frequency was fit with a generalized extreme value distribution with the location of the maximum being assigned as the preferred frequency (figure 6.3b).

Demarcating cortical areas A1, AAF, and SRAF

The axes for each cortical area were determined by first projecting a vector through the low and high frequency poles of each cortical area. Next, preferred frequency at each pixel intersecting this vector was plotted as a function of length along each vector. The edges of each axis were drawn at each reversal point, such as the high-low-high frequency reversal at the caudal border of A1. Lastly, these vectors were widened 5 times to accommodate additional pixels perpendicular and adjacent to the pixels comprising the vector. For comparisons across animals, these vectors were normalized and collapsed into 20 evenly spaced bins that concatenated pixels perpendicular to each bin.
**Measuring plasticity in the PTS zone**

To measure the cortical enhancement in the PTS zone, we computed the mean dF/Fo evoked by each respective frequency inside an ROI determined by the portions of each cortical axis as being > 16 kHz as measured prior to noise exposure. Tested intensities were at threshold, threshold +10, and threshold +20 dB, unique for each frequency but fixed across animals and recorded days. Thresholds for each frequency were the means of the whole window thresholds, rounded up to the next tested stimulation level. Whole window thresholds were determined based on visual inspection of spatial maps of evoked magnitude for each frequency-intensity combination.

**Figure 6.3: Behavioral evidence of acoustic trauma induced hypersensitivity at spared frequency.**

(A) Line plots showing logistical functions fitted to detection performance for 8 kHz tones of single mice (thin lines) and concatenated across mice (solid lines) for days pre (light blue) and immediately post-exposure (dark blue).

*Inset:* linear fits of **(B)** Mean +- SEM of fractional change in slopes relative to pre-exposure in logistical fits of performance for **left:** low-frequency (8 kHz) and **right:** high frequency (32 kHz) detection for sham and noise.
exposed mice. Change in slopes are plotted across days for noise and sham exposed mice. * denotes significance p<0.05 Wilcoxon rank-sum test **(C) left:** slopes of functions fitted to (left) ABR wave 1b and (right) behavior performance are concatenated across days post exposure, and the resulting fractional change relative to pre-exposure are plotted for each sham and noise exposed mouse. Significance (p>0.05 Wilcoxon rank sum test) between sham and exposed mice were only found for behavior. **Right:** same analysis except for 32 kHz tones. No significant differences between sham and exposed mice were found for both ABR and behavior data.

**RESULTS**

*A high-frequency acoustic exposure induces a lesion restricted to high frequency portion of the cochlea.*

The sensitivity of a healthy mouse cochlea ranges at least 4 to 64 kHz with maximum sensitivity near 16 kHz (see pre-exposure ABR wave I thresholds, figure 6.1b). We noticed that the most often protocols for acoustic trauma in mice are presented via octave band at 8-16 kHz (Liberman and Gao 1995; Maison and Liberman 2000) and that higher frequency exposures have not been attempted likely due to engineering challenges in building high-frequency acoustical systems. Our motivation to produce trauma at >16 kHz stems from the desire to generate a high-frequency hearing loss profile similar to that of age-related hearing loss in both mice and humans, and noise-induced high-frequency hearing loss in humans. Moreover, moving a hypothetical lesion edge further away from the low-frequency minimum in cochlear sensitivity. Providing enough topographical real estate could provide a greater opportunity to replicate the hallmark of competitive topographical remapping, which is the enhancement of topographical representation at the edge but not far away from the lesion. So, we developed a protocol that presents octave band (16-32 kHz) exposure for 2 hours at 103 dB SPL in a carefully designed and calibrated exposure chamber (for details, see methods). This exposure reliably induced permanent threshold
shifts (40-60 dB) in ABR wave 1b for all frequencies at and greater than 16 kHz (figure 6.1b; n = 36 ears, 18 mice), which reflect significant loss of high-frequency tone evoked activity primary afferent neurons. Similar to ABR wave 1b threshold shifts, we also found permanent threshold shifts in DPOAEs for frequencies at and greater than 16 kHz, which reflect loss of cochlear amplification by OHCs (Liberman et al. 2004). We wanted to ascertain ourselves of the pathology of cochlear damage because sometimes non-invasive measures can miss interesting causes underlying loss of peripheral function. First, we examined counts of type-1 afferent connections with IHCs via co-labeling of synaptic ribbon and corresponding postsynaptic glutamate receptors (figure 6.1d). Counts revealed significant synapse loss (>50 %) at and above 32 kHz, relative to sham exposed mice (figure 6.1f). Synapse loss was negligible at other cochlear frequencies at and below 22 kHz. Next, we examined the counts of OHCs (figure 6.1e) and found significant cell loss at 45 and 64 kHz (figure 6.1g), but negligible loss at other cochlear locations. Considering that DPOAE threshold shifts were significant at locations representing frequencies low as 16 and 22 kHz, we attempted to assess OHC stereocilia integrity, which has previously been shown to correlate with single unit threshold shifts in primary afferent (Liberman and Dodds 1984). We found a systematic disarray in the stereocillia in the first row of noise exposed OHCs at locations representing exposure frequency 32 kHz (figure 6.1e) and higher. Evidence of OHC damage at PTS regions 16 and 22.6 kHz were not conclusive but do not discount OHC damage. Taken together, consistent thresholds in both DPOAE and ABR wave 1b indicate a functional lesion across all high-frequency areas (>16 kHz) of the periphery and that threshold shifts are due to loss in OHC function as supported by histological markers of pathology. Importantly, our histological analysis and measures of cochlear function reveal no changes at the low-frequency encoding portions of the cochlea.
Operant tone detection confirms high frequency hearing loss.

ABR wave 1 thresholds alone are not enough to indicate loss of detection thresholds (Chambers et al. 2016a) which suggests caution in interpreting our peripheral measures. To confirm high-frequency hearing loss, we trained mice to detect tones encoded in the unaffected (low frequency, 8 kHz) and PTS (high frequency, 32 kHz) portions of the cochlea. Tone detection was assayed with a head-fixed operant behavior task (figure 6.2a); water deprived mice are trained in a go-no-go task and correct responses were rewarded with water. Tone bursts were initially presented at a high intensity associated with high detection performance and lowered until detection threshold was determined using the 2-down 1-up adaptive tracking method. This adaptive tracking paradigm was flanked with tone presentations at +/- 5 dB SPL (figure 6.2b), which complements our measures of threshold with measures of sound detection performance across the range from chance performance to perfect detection (figure 6.2c). In noise exposed mice, we found a very transient temporary threshold shift (1 day, ~20 dB), but no permanent threshold shifts for low frequency detection. For high frequency detection, we found an immediate threshold shift (~65 dB), followed by incomplete recovery to a new permanent threshold shift (~45 dB, figure 6.2d). This high-frequency hearing loss is consistent with both ABR wave 1b threshold shifts (figure 6.2) and, importantly, loss of OHC integrity because OHC amplification sets absolute sensitivity thresholds independent of remaining afferent connections.

Behavioral evidence of perceptual hypersensitivity near threshold, at frequencies spared from acoustic trauma.

Given the topographically restricted nature of the PTS region, we anticipate a central reorganization that leading to enhancement of intact frequencies adjacent to the PTS region (i.e.
11.3 or 8 kHz) that would lead to enhancements in behavioral detection 8 kHz. Detection thresholds at 8 kHz were unaffected by exposure, so we examined the dynamics of tone detection across sound intensity. Using the slopes of modeled behavioral functions, we found enhanced detection performance at 8 kHz following trauma in our noise but not sham exposed animals (figure 3c, p < 0.05 Wilcoxon rank-sum test). Notably, this enhanced tone detection performance was immediate and persistent over all days tested post exposure, relative to controls (figure 6.3b, p < 0.05 Wilcoxon rank-sum test for each day bin). Persistent enhancement in tone detection thresholds were also unable to be explained by any commensurate changes in ABR wave 1b growth (figure 6.3c, p = .817 Wilcoxon rank-sum test). We initially focused on low-frequency detection because peripheral measures of integrity and cochlear function were completely preserved and thus would be presented at intensities most relevant to animals, but we also wondered if there might be evidence of cochlear or central recruitment at higher frequencies and intensities. For 32 kHz tones, our measures of ABR 1b wave growth and sound detection performance failed to show any differences between exposed and sham mice (Δ slope, ABR wave 1b p = 0.817, ABR; tone detection p = 0.534 Wilcoxon rank-sum test). We are careful to make any conclusions about high-frequency sound encoding due to the low number ABR amplitudes to fit our slopes, but we find it reassuring that lack of change in slopes are comparable across both ABR wave 1b and sound detection at 32 kHz. For 8 kHz tone detection, we predicted and found an enhancement in tone detection at a frequency whose sensory coding appear unchanged at the cochlea and that this may reflect compensatory plasticity in central pathways.
Figure 6.4: Calcium imaging in awake head-fixed mice provide magnitude of tone-evoked activity and topographical organization of ACtx.

(A) Spatial distribution of dF/Fo evoked by tones (4-64 kHz) presented at 40 dB SPL are plotted for the baseline of a representative mouse. (B) Time-course of dF/Fo measured in ROI (see box in (A) and (D)). Scalebar: dF/Fo= 1%, time=500 ms. (C) the cumulative distribution of dF/Fo magnitudes evoked by tones across all tone frequencies presented at 40 dB SPL. CDF=0.65 is the threshold used to determine sound-responsivity. (D) left: the area of ACtx evoked by tones at 40 dB SPL, across 4-64 kH. right: the tone intensity thresholds for each pixel. (E) The frequency-response area measured in ROI (see box in (A) and (D)). top: the normalized dF/Fo magnitudes evoked across frequencies presented at intensities near threshold (pixel threshold to threshold +20 dB SPL) are plotted (bars). These magnitudes are reduced to a preferred frequency by fitting a generalized extreme-value distribution
(dashed line), the location of the peak is determined as the preferred frequency. (F) The map of preferred frequency across the imaging surface of ACtx.

**Tracking daily cortical representation of sound frequency across A1, AAF, and SRAF using chronic widefield calcium imaging.**

Enhanced tone detection at 8 kHz is likely a result of compensatory plasticity in the central auditory system, as suggested by lack of pathology or enhanced output in measures of cochlear integrity and function, respectively. We wondered both whether compensatory plasticity was present in our mouse model of HFHL and whether this plasticity parallel enhancements in low frequency tone detection. To track the day-by-day spatial dynamics of cortical reorganization, we used an approach previously developed to track the sound coding of excitatory neurons across the auditory cortex using widefield epifluorescence imaging of genetically encoded calcium indicator GCaMP6s across the auditory cortex (Romero et al. 2019). This method provides a reliable measure for the magnitude of tone-evoked activity (figure 6.3b), which are then used to produce frequency response areas (FRAs) for each pixel (figure 6.3e). Reducing the FRAs near the respective pixel threshold (figure 6.4c-d) reliably revealed cortical areas A1 and AAF, as indicated by the low-high-low reversal in preferred frequency, and SRAF indicated by its proximity ventral to A1 and AAF, and its defining low-frequency island (figure 6.4f). Preferred frequencies are at each pixel independently of the other, except for small (3x3 pixel) gaussian filter applied to remove spurious pixels for dF/Fo maps at each intensity-frequency combination.

We found our maps to be stable over 30 days in our sham exposed mice, consistent with our previous report (Romero et al. 2019).
Figure 6.5: High-frequency acoustic trauma induces rapid cortical re-organization of sound frequency

(A) Sequence of preferred frequency maps at representative days relative to acoustic trauma. Outside black line represents area of sound responsive pixels prior to acoustic trauma. Inside black lines indicate cortical axes A1, AAF, and SRAF. **(B)** The percent of non-responsive (NR) pixels before and immediately after exposure, along the length of axes A1 (**left**) AAF (**middle**) and SRAF (**right**). Colored insets indicate spatial arrangement of preferred frequency prior to acoustic trauma. **Top:** NR pixels in a representative noise-exposed mouse. **Bottom:** Mean ± SEM of NR pixels across exposed and sham mice. 

**C** Mean ± SEM of NR pixels across days relative to exposure in noise or sham exposed mice in A1 (top) and AAF (bottom). 

**D** Preferred frequency plotted along length of axis before (solid) and immediately after (dashed) exposure for A1 (**left**) AAF (**middle**) and SRAF (**right**). 

**Top:** Preferred frequency in a representative noise-exposed mouse. Red and blue stripes indicate low (8 kHz +/- 0.5 octave) and high (32 +/- 0.5 octave) frequency. 

**Middle:** Mean ± SEM of preferred frequency across noise exposed mice. 

**Bottom:** Mean ± SEM of preferred frequency across sham exposed mice. 

**E** Mean ± SEM of percent change in cortical area representing low (8 kHz +/- 0.5 octave) and high frequencies (32 +/- 0.5 octave) (see blue and red areas in (D, top)) in noise and sham exposed mice across day for A1 (top) AAF (middle) and SRAF (bottom).

**Immediate and wholesale cortical reorganization follow acoustic trauma.**

Leveraging our ability to track the spatial (across cortical areas) and temporal (day-by-day) changes in cortical responsiveness and sound frequency representation, we sought to determine whether plasticity in the central auditory system would parallel changes in sound detection of spared low frequencies. We focused our analyses on areas of the primary auditory cortex that exhibit the functional organization and dynamic range of sound processing, i.e. axes of tonotopy. These axes, A1, AAF, and SRAF, are demarcated using the frequency extrema, low and high, in BF maps derived from imaging sessions collected prior to noise exposure (Figure 6.5a, also see insets in 6.5b). The spatial location of these axes was registered across days using the images of the surface vasculature, which enable us to track at each pixel changes sound representation relative to pre-exposure. Immediately after noise exposure, we found 40% of the PTS zone (BF >
16 kHz, pre-exposure) in A1 and AAF became non-responsive to sound (Figure 6.5c, vs control p < 0.05 Wilcoxon rank-sum test). Within the period of 2-6 days post exposure, pixels in the PTS zone recover sensitivity to sound (figure 6.5a,c), in both A1 and AAF. Recovery of 16 kHz responses were significant at SRAF but not A1 and AAF (day 0-1 to day 3-6, p<0.05 paired Wilcoxon rank-sum test). The lack of short-term recovery for any of the PTS frequencies in A1 and AAF suggests recovery in the PTS zone was likely facilitated either by disinhibition or strengthening of existing synapses (Chen et al. 2011), or an outgrowth of new contacts from spared frequencies. The cortical reorganization of preferred frequency towards those below the PTS zone was swift and sustained (figure 6.5e).
Figure 6.6: Cortical reorganization in PTS zone is caused by active enhancement of frequencies at edge of lesion.

(A) Schematic depicting alternative mechanisms for cortical reorganization. Reorganization can take place in the absence of plasticity (i.e. residual tuning, middle column) or in the presence of plasticity (i.e. enhancement of responses to frequency areas at the edge of the PTS zone, right column) top: normalized amplitude of evoked activity from 0 (blacks) to 1 (white) middle: preferred frequency as a function of length along axis bottom: tuning
Figure 6.6 (Continued)
curves depicting magnitude of neural evoked by tones of respective frequency, plotted along the axis. (B) Scatter plot of mean tone evoked dF/Fo in the deafferented zone (BF > 16 kHz, pre-exposure) in individual noise (solid circles) and sham (open circles) exposed mice. Solid line denotes mean. Black solid line with arrows at top denote significance and direction of difference, respectively between tone evoked dF/Fo in noise vs sham exposed mice (p < 0.5, Wilcoxon rank-sum test). No significant differences were found in mean dF/Fo prior to exposure for all tone frequencies (p > 0.05, Wilcoxon rank sum test). (C) left: Mean +- SEM of dF/Fo in PTS zone evoked by 11.3 kHz across days post-exposure in noise and sham exposed mice right: Mean +- SEM of dF/Fo in deafferented zone evoked by 45.3 kHz across days post-exposure. (*) denotes significance between trauma and sham groups (p < 0.05 Wilcoxon rank-sum test).

Change in the spared frequency response strength measured from the deafferented zone.
We observed rapid and substantial changes in sound frequency representation, but the possibility remains that compensatory plasticity was not responsible for cortical reorganization of preferred frequency. This point is illustrated (Figure 6.6a, middle), where residual tuning of spared frequencies are able to explain wholesale shifts in preferred frequency. We predicted that compensatory plasticity was responsible for cortical reorganization and may parallel perceptual enhancements at spared low frequencies. We tested this hypothesis by tracking across day the magnitude of tone evoked activity in the PTS zone of A1, AAF, and SRAF (BF > 16 kHz, pre-exposure). As expected, we found an enhancement of tone evoked activity for frequencies at the edge of the PTS zone, 11.3 kHz for A1, AAF, and SRAF (figure 6.6b, p < 0.05 Wilcoxon rank-sum test). 11.3 is the frequency closest to the lesion boundary (16 kHz), but 8 kHz is the next closest frequency and may be considered as part of the lesion edge as we found significant enhancement for 8 kHz at A1 and SRAF (figure 6.6b, p<0.05 Wilcoxon rank-sum test), but not at AAF (p > 0.05 Wilcoxon rank-sum test). Interestingly, we also found an enhancement of tone
evoked activity in the preserved zone for 8 and 11.3 kHz in A1, and 11.3 kHz in SRAF (p<0.05 Wilcoxon rank-sum test). The enhancement of edge frequencies in the intact zone may indicate that disinhibitory circuits may be more global than simply restricted to the deafferented zone, but this would require a finer tuned demarcation of bottom-up projections from PTS regions. Here, we find that the mean enhancement of edge frequency is immediate and sustained following exposure, but we also noticed that time-course of recovery had a couple of distinguishing features. First, we observed an immediate increase in tone-evoked responses for A1 and SRAF across animals (p<0.05 paired Wilcoxon rank-sum test) but not for AAF. Following this initial enhancement, we observed a bounce-back towards pre-exposure levels at day 7-14 (figure 6.6c; vs sham exposed mice p>0.0.5 Wilcoxon rank sum test). Ultimately, the long-term result of noise exposure pushes towards enhancement of edge frequencies 8 and 11.3 kHz relative to sham exposed mice for all cortical areas A1, AAF, and SRAF (p<0.05 Wilcoxon rank-sum test).

Differences in the time course of plasticity between cortical areas may reflect differences in local microcircuits, short range horizontal projections, long range projections, or a combination of the above.

**DISCUSSION**

We tracked the recovery of mice following acoustic trauma using a novel high-frequency exposure that lesions the periphery in the range of high frequency sensitivity (>16 kHz) while the lower half of frequency sensitivity (<16 kHz) is preserved (figure 6.1). Following trauma, we showed that mice in an operant behavior task reported loss of and preservation of perceptual sensitivity to high and low frequencies, respectively (figure 6.2), confirming high-frequency hearing loss. Whereas peripheral measures of physiological integrity and cochlear function
suggest preservation of low-frequency coding, we found a paradoxical increase in perceptual growth for low frequency detection that was immediate (within 0-1 days) and sustained (> 30 days, Figure 6.3). In attempt to track the neural correlate of enhanced behavior, we used a method previously developed (Romero et al. 2019) to track the day-by-day representation of sound frequency across the topographical axes A1, AAF, and SRAF of the auditory cortex (Figure 6.4). We found that the high-frequency edge of the PTS zone in A1 and AAF became non-responsive to sound immediately after acoustic trauma but recovered within 2-5 days while SRAF remained responsive throughout (Figure 6.5) due to its sensitivity across a wide frequency range (Figure 6.6b). We found the representation of sound frequency along each of the measured axes immediately shifted towards the lower, preserved frequencies (Figure 6.5c). The possibility remained that shifts in sound frequency representation were a result of residual tuning, but we tested and demonstrated an enhancement of evoked responses to tone frequencies at the edge (e.g. 11.3 and 8 kHz) of the PTS zone (BF > 16 kHz) of A1, AAF, and SRAF (figure 6.6b). Surprisingly, we also found an enhancement also in the preserved zone (BF < 16 kHz) for A1 and SRAF. The time course of edge frequency enhancement was immediate (days 0-2) and persistent (>30 days) parallel our measured enhancements in tone detection. The unique timecourse in the magnitude of cortical enhancement revealed among A1, AAF, and SRAF reveals clear differences in cortical plasticity in time and space (figure 6.6c).

Sound processing in the auditory cortex is dramatically modified following peripheral damage whether it be from NIHL (Norena and Eggermont 2005, Yang et al 2011, Resnik and Polley 2017, Auerbach 2019), due to ototoxicity (Chambers et al. 2016a, Resnik and Polley 2017, Takasu and Tateno 2019), or genetic factors. These different modes of hearing loss have significant impact the spatial and time course of recovery. Remarkably, we demonstrated a rapid
recovery (~2 days) in responsivity to sound in the PTS zone following topographically restricted, bilateral lesions which were much slower compared to the slower recovery in firing rates following broadband, unilateral and bilateral nerve loss (7-30 days), but consistent with time-course in single neuron receptive field remapping following temporary threshold shifts (Resnik and Polley 2017). The restricted lesion model of hearing loss presented in this study provides a unique opportunity to observe the facilitation by intact sensory areas in the recovery in the PTS zone. Additionally, the two-octave range of intact sensitivity provides internal control within each mouse for contrasting against any identified mechanistic correlate of cortical reorganization, such as a differential activation or disinhibition of unique inhibitory circuits within the cortex (Yang et al. 2011).

Reining in dramatic increases of excitatory circuits are the responsibility of inhibitory circuits. These circuits comprise 3 general classes of interneurons, parvalbumin (PV)-expressing, somatostatin (SST)-expressing, and vasoactive intestinal peptide (VIP)-expressing. Pyramidal neuron activity is shaped by unique patterns of innervation from these interneurons, PV target perisomatic regions (Freund and Katona, 2007; Tremblay et al. 2016) and SST target distal dendrites (Wang et al. 2004; Gentet et al. 2012). In turn, both PV and SST are inhibited by VIP interneurons. Recent work has focused on the unique role each of these interneurons have in shaping the activity of excitatory neurons. In one study of the somatosensory system, higher order thalamic inputs were responsible for suppressing SST while enhancing PV and VIP interneurons, a dynamic which was implicated in gating long term potentiation (Williams and Holtmaat 2019). In another study, the circuit of these three interneurons are implicated in the hyperactivity of pyramidal neurons in a neuropathic pain model induced by peripheral lesions; promisingly they found that enhancement of SST interneurons reduced hyperactivity and
reversed the neuropathic pain (Chichon et al. 2017). These studies highlight the unique importance each interneuron may have on the homeostatic and hebbian circuits that are controlled by higher order inputs and those circuits that may result in pathological elevations in sensory evoked firing rate.

A hallmark of compensatory plasticity in the auditory system is the graded recovery of firing rate along ascending stations along the ascending lemniscal pathways of the auditory system (Salvi et al. 1990; Auerbach et al. 2014; Chambers et al. 2016a; Chambers et al. 2016b). Similarly, the completeness of topographical reorganization is graded along the lemniscal pathway (Robertson and Irvine 1989; Kamke et al. 2003; Irvine et al. 2003). These observations lead us to the natural question, where does compensatory plasticity originate? We have a few clues. First, lesions of the somatosensory cortex suggest that related structures are required for the induction but not maintenance of plasticity relating to restricted peripheral lesions in the somatosensory thalamus (Parker and Dostrovsky 1999). Second, unlike the auditory and somatosensory system, topographical plasticity is incomplete or non-existent in the visual thalamus and divergence of thalamic projections are insufficient to explain cortical recovery following bilateral retinal lesions; yet topographical reorganization persists in the visual cortex and likely arises due to new or strengthened horizontal cortico-cortico projections (Gilbert and Wiesel 1992). Third, topographical projections from the lemniscal auditory thalamus appear to be fixed in adults (Barkat et al. 2011; Takesian et al. 2018) suggesting that the topographical nature of cortical plasticity in the auditory system is facilitated locally within discrete brain areas and not due to changes in the organization of ascending projections between areas. The coupling of perceptual enhancement in low frequency detection and the enhancement of representation of the same
frequencies in the auditory cortex provide evidence for a significant role of the auditory cortex in facilitating compensatory plasticity.

Taken together, our results suggest that enhanced tone detection following sudden high-frequency hearing loss reflects cortical reorganization and enhancement of frequencies at the edge of the lesion. This reorganization occurs across A1, AAF, and SRAF, albeit to different degrees which might reflect differences in local circuits, short-range horizontal, or long-range projections to the respective areas. Governing the dynamics of this competitive plasticity requires a precise sequence of disinhibition followed by E-I rebalancing circuit that is coordinated across spatial scales. This circuit may present both a blessing and a curse. On one side of the coin, this circuit may be particularly vulnerable to pathological forms of hyperactivity (e.g. Chicon et al. 2017) or runaway excitation. On the flip side, this circuit may present a powerful target for cortical remapping-based therapies for perceptual disorders related to hearing loss.
CHAPTER 7

CONCLUSIONS

The work completed in this dissertation ranged from sensory transduction at the periphery to behavior decisions, while traveling through many of the auditory structures that connect the two. While each of these structures (e.g. cochlea, cochlear nucleus, inferior colliculus, auditory cortex) rightfully deserve dedicated study such that careers spent doing so already have and will continue to provide breakthrough advances in our general scientific understanding, the approach taken in this dissertation was appropriate and revealing. The rudimentary modes of sensory processing at the early stages along the auditory pathway are transformed into become more abstracted modes of processing at later stages. Exploring these transformations towards abstracted modes of processing will be interesting using animal models with healthy hearing. On the flip side, using animal models with dramatically altered modes of stimulation, whether they be due to peripheral trauma or due to obstinately modified stimulation via neuroprostheses, could be particularly useful approaches for exploring sensory processing, as demonstrated in this report. Moreover, approaches that focus on the abstracted modes of sensory processing at later stages of the auditory pathway could provide a new chapter in the implementation of sensory restoration devices.
REFERENCES


individual cochlear implant fitting. European Archives of Oto-Rhino-Laryngology, 265(11), 1435.


