Spectral cues are necessary for encoding the azimuthal map of auditory space in the mouse superior colliculus

Shinya Ito[1*], Yufei Si[2], David A. Feldheim[2], Alan M. Litke[1*]

1. Santa Cruz Institute for Particle Physics, University of California, Santa Cruz
2. The Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz

Abstract

Sound localization plays a critical role in animal survival. To compute the incident sound direction, the animal can use three cues: interaural timing differences (ITDs), interaural level differences (ILDs) and the direction-dependent spectral filtering of the sound by the head and pinnae (spectral cues). Compared to ITDs and ILDs, little is known about how spectral cues contribute to the neural encoding of auditory space. Here we report on auditory space encoding in the mouse superior colliculus (SC). We show that the mouse SC contains neurons with spatially-restricted receptive fields (RFs) that form a topographic map of azimuthal auditory space. By eliminating each sound localization cue from the stimuli, we found that nasal RFs require spectral cues and temporal RFs require ILDs. Therefore, the lack of either cue results in the disruption of the azimuthal topographic map. These results demonstrate an unexpected role of spectral cues in azimuthal sound localization.

Introduction

Sound localization is the result of the sophisticated analysis of vibrations detected by an animal's ears. Unlike vision or touch, where the receptor position along the sensory detector encodes spatial information, the incident direction of the sound source must be computed. This calculation is based on three cues: interaural timing differences (ITDs), interaural level differences (ILDs) and the spectral modification of the sound as it enters the ear (spectral cues). Since Lord Rayleigh's seminal finding that ITDs and ILDs are used for low-frequency and high-frequency sound localization, respectively (the duplex theory)[1], ITDs and ILDs have been considered to be the two primary cues used for sound localization in the horizontal plane (except for small mammals that do not use ITDs[2–4]). On the other hand, spectral cues were hypothesized to be used for resolving directions that the ITDs and ILDs cannot distinguish,

namely, for resolving the front–back ambiguity and for determining the sound source elevation[3]. However, it is clear that spectral cues can also be used for horizontal-plane sound localization because ferrets and blind humans can be trained to perform a horizontal sound localization task with only one ear (i.e. without ITDs or ILDs)[5,6] leading to questions about the validity of the duplex theory of sound localization[3]. Therefore, we hypothesized that spectral cues also contribute to azimuthal sound localization.

To determine the mechanisms used to compute sound source direction, we conducted an electrophysiological study in the mouse superior colliculus (SC, also known as the optic tectum (OT) in non-mammalian species). The SC is an ideal brain area to study sound localization because it contains spatially tuned auditory neurons and a topographic map of auditory space. The properties of the SC/OT auditory map have been reported in a variety of species including barn owls[7,8], cats[9], ferrets[10], and Guinea pigs[11,12]. The SC also contains maps of visual and somatosensory space and is a well-studied area for topographic map formation[13]. Interestingly, a topographic map of auditory space has not been demonstrated in the mouse SC. In the present study, we used the head-related transfer functions (HRTFs) to present virtual auditory space (VAS) stimuli[2,11,14] to an alert, head-fixed mouse while recording from neurons in the SC. This allowed us to measure the receptive field (RF) properties of auditory neurons, determine the topography of the auditory spatial map, and measure the relative contribution of ITDs, ILDs, and spectral cues to the encoding of a spatial map. We found that the mouse SC contains a topographic map of auditory space along the azimuthal axis; spectral cues and ILDs are used to compute localized auditory RFs, but ITDs are not; and the relative importance of spectral cues and ILDs depends on the azimuthal position of the RF. Based on these findings we developed a mathematical model that recapitulates our data. These results demonstrate an unexpectedly important role for spectral cues in the formation of the azimuthal sound map, lending new insights into how mice perform sound localization, and demonstrate that the mouse is a useful model to study the mechanisms of auditory processing.

Results

Mouse SC neurons have spatially localized auditory RFs that are topographically organized

To determine the auditory response properties and organizational features of mouse SC neurons we used VAS stimuli and large-scale silicon probe recordings (Fig. 1a). We first measured a set of HRTFs, the modulation of sound by the physical structure of the head and pinnae, that contains each of the three sound localization cues: ITDs, ILDs, and spectral cues (Fig. S1). We used the HRTFs to filter a 100-ms white noise burst stimulus and played it via calibrated earphones so that the mouse would hear the sound as if it came from the direction represented by the corresponding HRTF (see Methods). The stimuli were presented at grid locations of 17 azimuths (−144° to 144°) and 5 elevations (0° to 80°), totaling 85 points in a two-dimensional directional field.

To measure the response properties of the SC neurons, we performed electrophysiology using 256-channel multi-shank silicon probes (Fig. 1b) while the VAS stimuli were presented to mice that were awake and allowed to locomote on a cylindrical treadmill (Fig. 1a). In 20 recordings from 10 mice, we recorded the spiking activity of 3556 neurons (excluding axonal signals; Fig. S2) from the deep SC (300–1600 μm from the surface of the SC; Fig. 1b, c). The neurons had a variety of temporal response patterns (Fig. S3a–d), with their peak response time having a bimodal distribution. Because most (77.5 ± 0.6%) neurons had a peak response faster than 20 ms after the stimulus onset (Fig. S3e), we used the spikes with latencies less than 20 ms for our analysis (see Fig. S3 for further discussion). We observed significant ($p < 0.001$) auditory responses from 56.7 ± 0.8% (n = 2016) of all the neurons identified (see "Significance test for the auditory responses" in Methods). To determine the RF properties of these neurons, we parameterized the RFs by fitting the Kent distribution[15] to the directional auditory responses of these neurons. The RFs were considered significant if the Kent distribution explained the data better than a flat distribution (i.e. Bayesian Information Criterion of the Kent distribution is smaller than that of a flat distribution; see Methods for details). Of the auditory responsive neurons, 22.8 ± 0.9% (n = 459) had a significant RF tuned to a specific sound source direction (examples: Fig. 1d–f).

To determine if the mouse SC contains a topographic map of auditory space, we tested if the location of the RF azimuth and the anterior–posterior (A–P) or medial–lateral (M–L) position of the neurons in the SC are correlated. We found that the SC neurons have a continuous distribution of preferred azimuths in the horizontal plane (Fig. 1g), with the RF azimuth linearly related to the neurons' A–P location (Fig. 1h). Using line fitting to measure the slopes (considering estimated systematic errors, see Methods) we calculated the slope of the auditory map to be 58 ± 4 °/mm. This slope is ~20% smaller than the slope measured for the visual map, as calculated by multi-unit activity in the superficial SC (73 ± 5 °/mm). We also found that the size of a neuron's RF increases as a function of the azimuth (Fig. 1i). The slope calculated between the RF elevations and M–L positions (13 ± 5 °/mm) is not as steep as the slope along the azimuth (Fig. S4). Comparing the response properties of each neuron when the mouse was stationary vs. running revealed that locomotion increased the spontaneous firing rate and reduced the stimulus-evoked firing rate, but had no influence on the properties of the topographic map (Fig. S5). These results demonstrate, for the first time, that the mouse SC contains a topographic map of auditory space along the A–P axis of the SC.

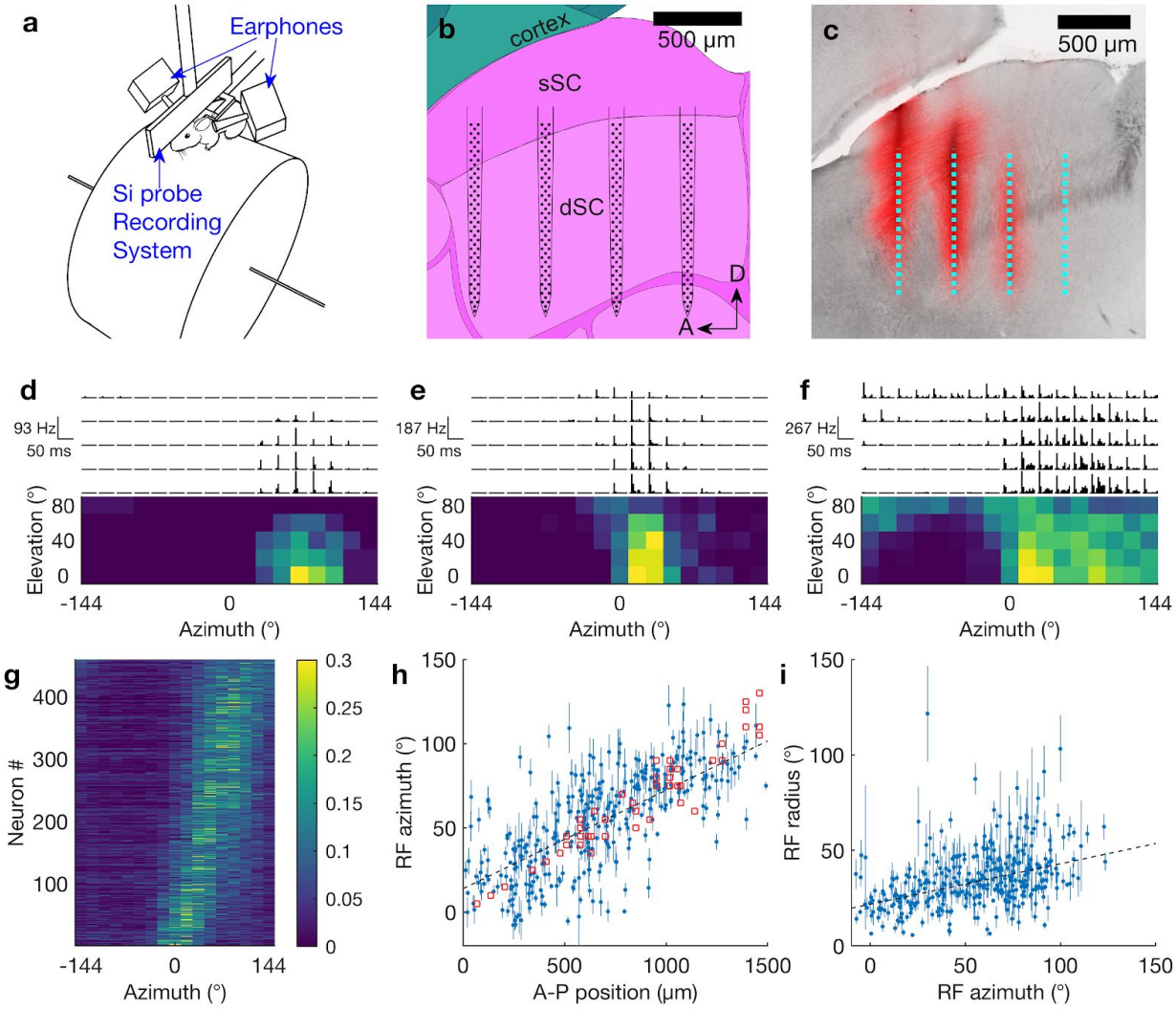

Figure 1: Auditory receptive fields (RFs) of SC neurons and their topographic organization.

a: An illustration of the virtual auditory space (VAS) stimulus and electrophysiology recording setup. During the recording, the head-fixed mouse can run freely on a cylindrical treadmill, while motion is recorded by a rotary encoder. Auditory stimuli are delivered through a pair of earphones near the mouse's ears. The entire setup is enclosed in an anechoic chamber.

b: A schematic of a 256-electrode silicon probe in the SC (sagittal plane). sSC: superficial SC; dSC: deep SC; A: anterior; D: dorsal.

c: A photograph of a sagittal section of the SC and the fluorescent tracing of DiI (red channel) that was coated on the back of the probe shanks. The dotted cyan lines indicate the active areas of the probe during the recording, reconstructed from the DiI traces (the location of the rightmost shank was inferred by a DiI trace in the adjacent slice). The four shanks with a 400-µm pitch nicely match the anterior-posterior extent of the SC.

d–f: Examples of the spiking activity and localized RFs of three representative mouse SC neurons. The top panels are the post-stimulus time histograms (PSTHs) in response to 85 virtual sound source locations (5 × 17 elevation-azimuth grid). The bin size of the histogram is 5 ms; the vertical range is from 0 to the maximum firing rate for the plotted neuron. The bottom panels are a summary heatmap of the activity within the 5–20 ms time interval of the PSTHs in the top panel. Brighter color indicates a higher firing rate. The RF is visible as a bright area of the figure.

g: Responses of all the neurons that have a localized RF to a stimulus from 17 horizontal directions. For each neuron, the scale of the responses is normalized by dividing by the sum of responses across all directions. Neurons are sorted by their tuned azimuth.
h: A scatter plot of anteroposterior (A–P) SC positions vs. RF azimuths showing topographic organization along the A–P axis of the SC. Each blue dot represents the auditory RF azimuth and the A–P position of an individual neuron (the error bars represent the statistical errors; it does not include the systematic errors discussed in Methods); red squares are visual RFs measured by multiunit activity in the superficial SC. The slope of the auditory RFs is 58 ± 4 °/mm (measured including systematic errors; see Methods for details) and that of the visual RFs is 73 ± 5 °/mm. The offset of the auditory RFs was 14 ± 3°. There is a strong correlation between the auditory RF position and A–P position within the SC (r = 0.70; Pearson's correlation coefficient).
i: A scatter plot of RF azimuth vs. RF radius of individual neurons. A positive correlation (slope: 0.21 ± 0.02; r = 0.41) of the RF azimuth and the estimated RF radius (specified by the concentration parameter κ of the Kent distribution (see Methods)) was observed.

Eliminating spectral cues results in the largest change to the RFs of the SC neurons

To determine the contribution of each sound localization cue to the formation of spatially restricted RFs, we presented the mouse with stimuli that lack a specific sound localization cue and observed the corresponding change, if any, to the RF structure of each neuron with and without cue elimination. The stimuli include the original stimulus that contains all cues (played twice as a control for assessing reproducibility of the RFs) as well as stimuli that lack ITDs, ILDs, or spectral cues. In this experiment, we recorded from 991 neurons from 6 mice, and found that 59.7 ± 1.6% (n = 592) had a significant auditory response and 16.4 ± 1.5% (n = 97) of these had a localized RF. We found that elimination of ITDs had no effect on the RF structures, consistent with the idea that the distance between the ears of the mouse (~1 cm), is too small to detect ITDs (~29 μs). However, when ILDs or spectral cues were eliminated, the SC neurons changed their RFs; some were affected only by spectral cue elimination (Fig. 2a–e), while others were affected only by ILD elimination (Fig. 2f–j).

To quantify the change of the RFs induced by each cue elimination, we used a cosine similarity index (SI). Namely, the response firing rates at 85 spatial locations are considered as a vector, and a normalized inner-product was calculated between the responses to the original and the cue-eliminated stimuli (see Methods for details). If a neuron has the same RF structure with the cue-eliminated stimulus, the SI is 1; if the RF structure is uncorrelated with the original, the SI is 0. The SI of the responses to the control and the original stimuli measures the reproducibility of the RF. Of the neurons with spatially localized RFs, 53 ± 5% (n = 51) had a 'reproducible' RF (that is, the SI of the control dataset was significantly ($p < 0.01$) larger than 0), and only these neurons were used in the following analysis. The SI value of the control dataset was used to normalize (by simple division) the SI values of the cue-elimination results (normalized SI: NSI). The loss of NSI (1 − NSI) shows the dependency of an RF to a particular sound localization cue.

The horizontal RFs of all the neurons with a reproducible RF are shown in Fig. 2k–o. Overall, elimination of spectral cues caused a drastic change to the RFs leading to the largest loss of NSI, followed by the ILD elimination, while the ITD elimination did not cause a significant

loss of NSI (Fig. 2p). These results highlight the importance of spectral cues for the calculation of auditory RFs in the SC.

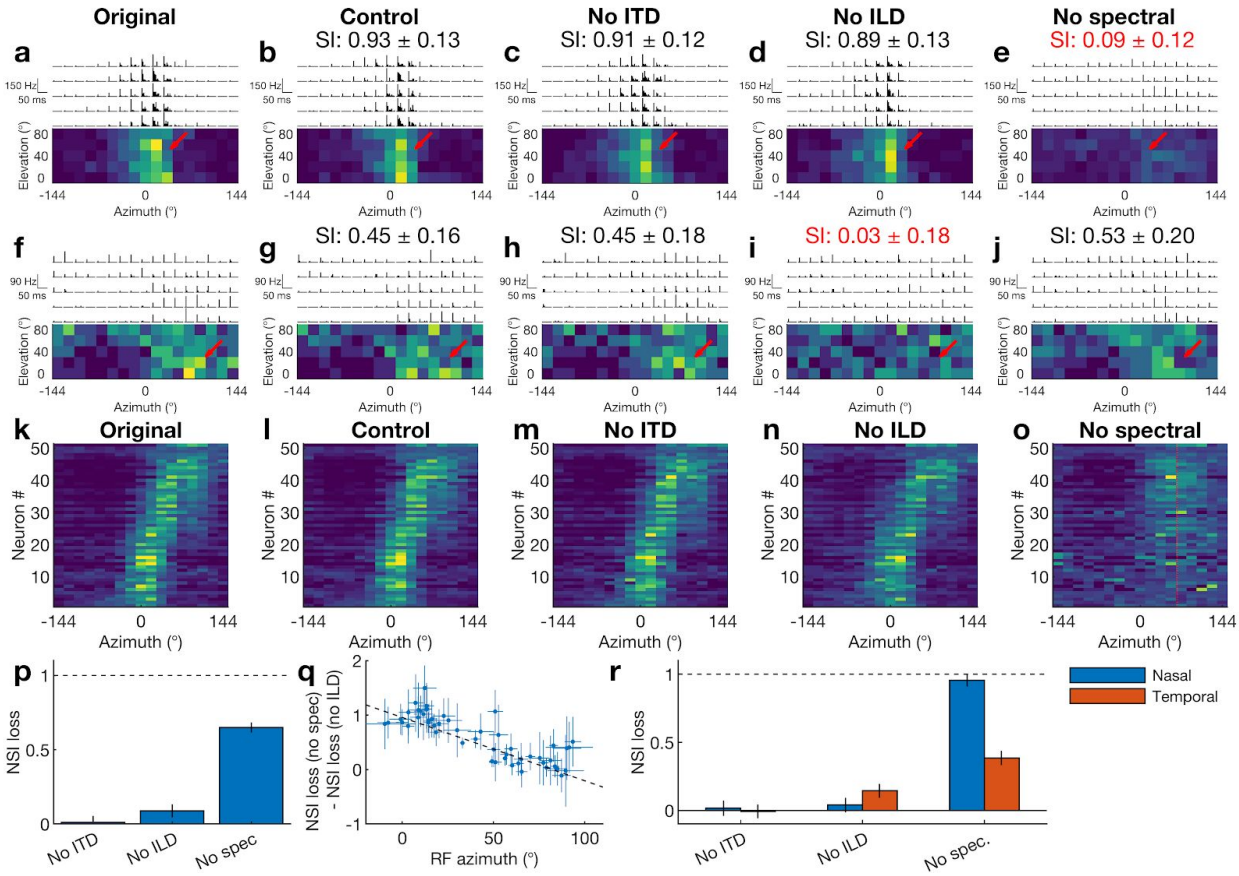

Figure 2: Structural changes of auditory RFs due to ITD, ILD, and spectral cue elimination.

a–e: RFs of a single neuron in response to the original (a), control (b), ITD elimination (c), ILD elimination (d), and spectral cue elimination (e) stimuli. Fig. b–e indicate the similarity index (SI; see text). This neuron's RF structure did not change during ITD elimination and ILD elimination, but it disappears when spectral cues are eliminated (red arrow).

f–j: Example responses from another neuron whose RF disappears with ILD elimination (red arrow).

k–o: Summary of the Horizontal RFs of all 51 neurons that had a reproducible RF. For each neuron, the scale of the responses is normalized by dividing by the sum of responses across all directions. The range of the color scale is from 0 to 0.3. The neurons are sorted by their tuned azimuth. Little changes in the RFs are observed with ITD and ILD elimination (m, n), but a dramatic change of the RF structure was observed with spectral cue elimination (o). With spectral cue elimination, all the RF centers become close to 65° azimuth where the ILD is maximized (red dotted line).

p: Summary of NSI losses in the cue elimination experiments. A large NSI loss is observed with spectral cue elimination, consistent with (m–o).

q: A scatter plot of the RF azimuths and the differences of spectral-cue and ILD NSI losses. As the RF azimuth increases, neurons depend less on spectral cues and more on ILDs. ($r = -0.83$).

r: Comparison of the average NSI loss for nasal and temporal RFs. Neurons with nasal rather than temporal RFs depend almost entirely on spectral cues.

Heterogeneous use of spectral cues and ILDs across the SC

The cue elimination experiment also revealed that neurons that rely on spectral cues and ILDs to compute RFs are not distributed evenly across the SC. Neurons with nasal RFs (in the anterior SC) have a stronger dependence on spectral cues, while neurons with temporal RFs (in the posterior SC) have a stronger dependence on ILDs (Fig. 2q). When the neurons were divided into nasal and temporal groups at 45° azimuth, we found that the neurons in the nasal group relied almost completely on spectral cues to form their RFs, while those in the temporal group relied on both ILDs and spectral cues to form their RFs (Fig. 2r). As a consequence, when spectral cues are eliminated from the stimulus, most of the nasal RFs disappear and those that remain have a peak firing rate at ~65° azimuth (Fig. 2o, red dotted line), where the ILD is at its maximum (Fig. S1e). This shows that spectral cues are a necessary component for computing an azimuthal topographic map in the SC.

Responses of SC neurons to monaural or extended ILD stimuli confirm the importance of spectral cues for encoding the azimuthal topographic map

We were surprised to find that ILD elimination did not have a more dramatic effect on the RFs of SC neurons. One reason for this could be that we only corrected the average ILD across the frequency range of 5–80 kHz. Because ILDs are a function of frequency, and can be as high as ±30 dB in a high-frequency range (Fig S1c, d), the ILD elimination stimulus may have had a remaining ILD in the high-frequency range that can be utilized by neurons (as suggested in a review study[16]). Therefore, we conducted two additional experiments to clarify the role of ILDs in RF formation. In one, we presented monaural stimuli, which entirely lack physiologically relevant ILDs and spectral cues from the ipsilateral ear. In the other, we extended the ILD stimuli to include a range of ±40 dB ILD.

When we compared the RFs of neurons generated from binaural (control) and monaural stimuli with and without spectral cues, we found that monaural stimuli led to a loss of temporal (Fig. 3i, m), but not nasal RFs (Fig. 3d, i, m). An example neuron with a nasal RF is shown in Fig. 3a–e. Note that spectral cue elimination, but not monaural stimuli, led to a loss of SI. Fig. 3f–j show a neuron that has an RF that is altered both by spectral cue elimination and monaural stimuli. In this neuron, the response to the sound coming from the front was not changed during monaural stimuli (blue arrows), but the difference in the responses to ipsilateral and contralateral stimuli goes away during monaural stimuli (Fig 3i, j, red arrows). This trend holds for every neuron with a localized RF (Fig. 3k–m), suggesting that the spectral cues are responsible for nasal RFs, and ILDs are responsible for temporal RFs. The overall NSI loss for each elimination was similar, but they act on different parts of the RFs; hence, elimination of both cues leads to the complete loss of NSI (Fig. 3n).

We also used extended ILD stimuli to see if neurons are tuned to a specific value of the ILD or if the firing rate increases monotonically with the ILD. We kept the spectrum of the

stimulus flat, but changed ILDs randomly across the range from −40 dB to +40 dB (see Methods). We recorded from 856 neurons from 5 mice, and found that 9.2 ± 1.0% (n = 79) of the neurons had significant and non-flat auditory responses across the range of ILDs tested (Fig 3o). Only one neuron displayed a peaked tuning (a peak firing rate that appears anywhere between −40 dB and +40 dB, and is significantly ($p < 0.01$ after Bonferroni correction) higher than the firing rate at both ±40 dB). This means that the firing rate of most neurons increased monotonically as a function of the ILD (Fig. 3o), supporting the observations in Figs. 2o and 3l that spectral cue elimination makes the RF of ILD-dependent neurons shift to where the ILD is maximized (at ~65° azimuth; red dotted line in Fig. 2o). These results suggest that the neurons cannot form a topographic map by ILDs alone.

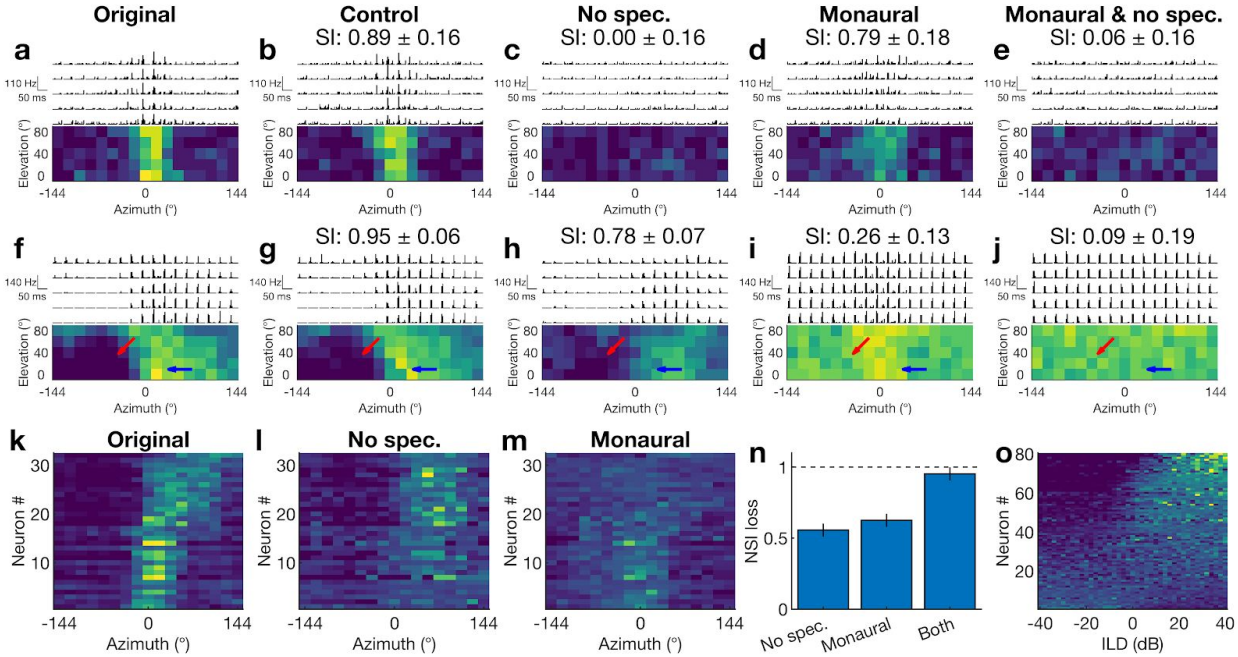

Figure 3: Structural changes of auditory RFs due to spectral cue elimination and monaural stimuli.
a–e: Response properties and RFs of a neuron in response to binaural original (a) and control (b) stimuli, binaural stimuli with no spectral cues (c), monaural stimuli (d) and monaural stimuli that also lack spectral cues (e). This neuron loses its RF with spectral cue elimination (c, e), but not with monaural stimuli (d).
f–j: An example of a neuron whose RF is comprised of spectral cue and ILD components. Notice how the temporal part of the RF remains intact during spectral cue elimination, suggesting that this part is based on ILDs. The nasal part of the RF remains intact with monaural stimuli, suggesting that this part of the RF is based on monaural spectral cues. With the monaural stimulus, the ipsilateral side has responses (red arrow) that are not seen in binaural stimuli (f–h).
k–m: Horizontal RFs in response to the original binaural stimuli, and binaural stimuli with no spectral cues and monaural stimuli. For each neuron, the scale of the responses is normalized by dividing by the sum of responses across all directions. The range of the color scale is from 0 to 0.3. When spectral cues are eliminated, most of the nasal RFs are distorted, and the temporal RFs are maintained (l). When monaural stimuli are presented, the nasal RFs are maintained, the temporal RFs are eliminated, and suppression of the ipsilateral sound is reduced.
n: Summary of the NSI loss to the stimuli with no spectral cues, monaural stimuli and both. The stimuli with no spectral cues and monaural stimuli have a similar level of NSI loss, but their roles are different. When both are eliminated, the NSI is close to 1, meaning that the RFs are completely lost.

o: Responses of the ILD sensitive SC neurons to an extended range of ILDs. The firing rates (FRs) are normalized by the FR sum across the full ILD range for each neuron. Neurons with non-flat responses to ILDs (a flat distribution fit gives a large χ^2 value ($p < 0.01$)) are shown. Most of the neurons display a monotonic change of the FR as a function of the ILD, and are not 'tuned' to a specific ILD value. The range of the color scale is from 0 to 0.12.

A mathematical nasal vs. temporal RF model describes the components used to form spatially restricted RFs and predicts the results of the cue-elimination experiments

Given that spectral cues and ILDs contribute to different azimuthal parts of the RF, we developed a mathematical model that explains how spatially restricted RFs form along the azimuthal axis. Our model is quite simple, consisting of only two basis RFs and 4 parameters for each neuron; we call it the nasal vs. temporal RF model. The goal of the model is to describe most of the observed neuronal RFs using simple basis RFs that correspond to those made by spectral cues and ILDs. The parameters are the baseline firing rate (b), the weight and the concentration (that specifies the RF radius) for the spectral-cue-based 'nasal' RF (w_1, κ), and the weight for the ILD-based 'temporal' RF (w_2) (Fig. 4a; see Methods for the model details). Notice that the nasal RF (an even function of the azimuth) and the temporal RF (an odd function) act as two orthogonal basis functions. Hence, despite the small number of parameters, the model can efficiently express different RF positions across the horizontal plane simply by changing the ratio of w_1 and w_2. Indeed, the model fits well to 82 ± 2% of the neurons with the localized RFs in Fig. 1g when fit to the horizontal RF (elevation = 0°), and, even though the model does not have a component that distinguishes elevations, it still fits well to 53 ± 2% of the neurons when fit to full-field RF (including all elevations) (Fig. 4b).

The model also reproduces the heterogeneity of how ILDs and spectral cues are used in the SC. The ratio of the weights w_1 (weight on spectral cue) and w_2 (weight on ILDs) changes as a function of the RF azimuth (Fig. 4c). This is consistent with the results in Fig. 2q, which show that neurons with nasal RFs depend on spectral cues and neurons with temporal RFs depend on ILDs.

This nasal vs. temporal model also predicts the results of the cue-elimination experiments. Even though the model parameters are achieved by fitting only to the 'original' data of the cue-elimination experiment, by erasing the ILD-component (by setting w_2 = 0) or the spectral component (w_1 = 0), it qualitatively reproduces the disappearance of the RFs in the cue-elimination experiment (Fig. 4d, e; compare with Fig. 2a–j). This model also predicts the changes to the RFs of neurons during the monaural stimulation experiments (assuming that monaural stimuli have a large ILD (i.e. ILD-RF = 1) at every virtual location, Fig. 4f, g; compare with Fig. 3a–j).

Despite its simplicity, the nasal vs. temporal RF model recapitulates our findings. Namely, the RF azimuths are tuned by changing the ratio between the spectral-cue-based nasal RF and the ILD-based temporal RF. Therefore, the RF position infers relative dependence on spectral cues and ILDs that changes across the anteroposterior axis of the SC, and the complete elimination of either cue leads to the disappearance of a topographic map.

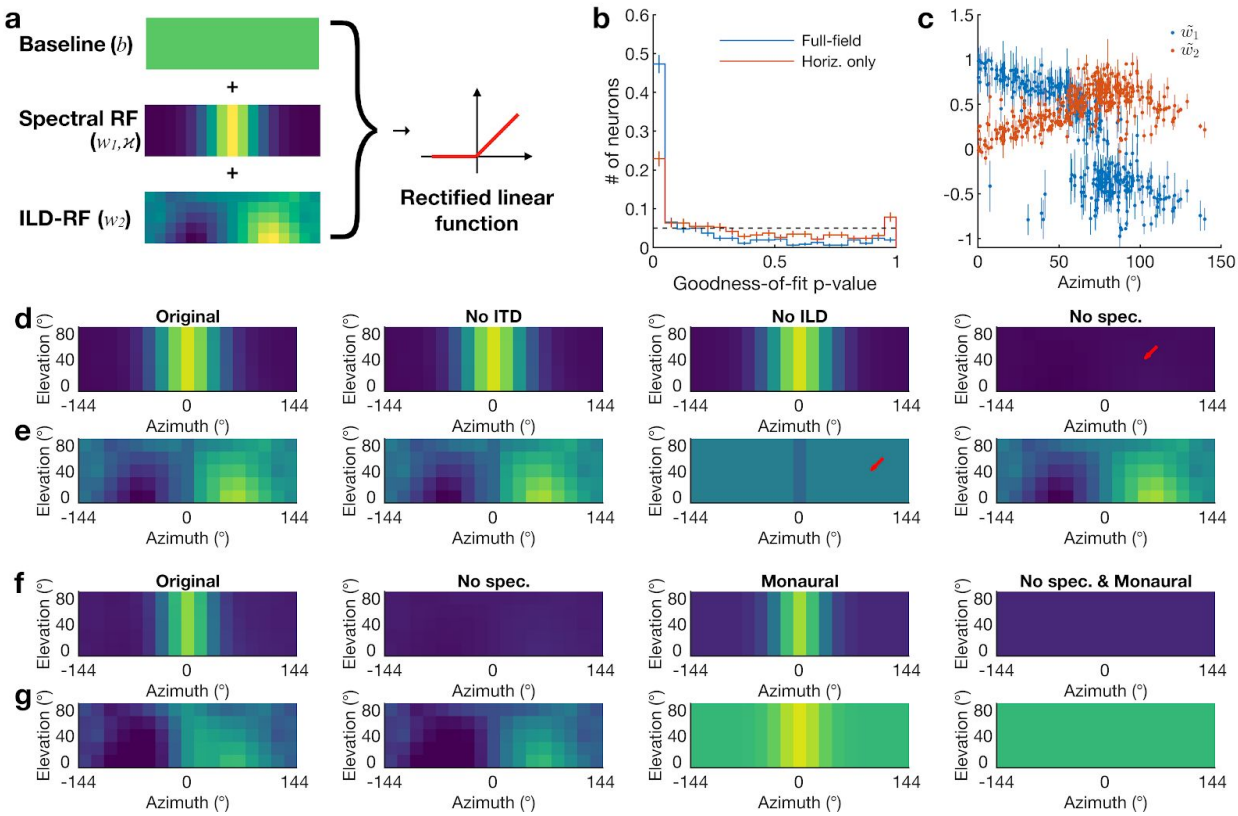

Figure 4: Modeling SC auditory RFs.

a: A schematic of our model of SC neurons' RFs. The model consists of a linear addition of the baseline (b), the spectral-cue-based RF (w_1 , κ), and the ILD-based RF (w_2), with a rectified linear function applied at the end (see Methods for the model details).

b: Goodness-of-fit p-value distribution of the model of the observed RFs. The black dashed line represents the expected distribution when the model perfectly describes the data (flat distribution). Our model fits well to 53 ± 2% of the neurons (below the dashed line) for the full-field stimuli and 82 ± 2% for the horizontal plane stimuli.

c: Normalized parameters of the fit as a function of the RF azimuth. Normalization is performed by dividing the parameters by the sum of the absolute values of w_1 and w_2 . Consistent with Fig. 2q, the dependence on spectral cue (w_1) is larger for the nasal RFs, and the dependence on ILD (w_2) is larger for the temporal RFs.

d, e: RF changes in the cue-elimination experiments predicted by the model. The fit was performed on the RFs shown in Fig. 2a, f, respectively. The model predicts the disappearance of the RF due to spectral cue elimination (d) and ILD elimination (e) (red arrows; compare with Fig. 2a–j).

f, g: RF changes in the spectral cue elimination and monaural stimuli experiments predicted by the model. The fit parameters are based on the RFs shown in Fig. 3a, f, respectively. The model predicts the disappearance of the RF by spectral cues (f), as well as the more complex behavior of the neuron in Fig. 3f–j (g).

Discussion

In the present study we: (1) determined that the mouse SC has a topographic map of azimuthal auditory space; (2) discovered that SC neurons rely on both spectral cues and ILDs to create spatially restricted RFs and form this topographic map; (3) found that the contribution of

each of these cues varies across the azimuthal direction; and (4) developed a simple mathematical model that summarizes these findings.

The mouse has a topographic map of azimuthal auditory space in the SC

By recording from SC neurons in awake-behaving mice in response to VAS stimulation, we found that the mouse SC contains auditory neurons that form an azimuthal topographic map of sound, with the slope of azimuth approaching that of the visual map in the SC. Specifically, we found a strong correlation between the A–P position of the auditory neurons in the SC and their RF azimuths (Fig. 1h, r = 0.70). The correlation of the M–L neuron position and RF elevation was weaker (Fig. S4, r = 0.22). This is consistent with a previous report that showed that mice can discriminate two sound sources better along the azimuthal axis than the elevation axis (azimuthal: 31 ± 6°, elevation: 80.7 ± 1.7°)[17].

An auditory topographic map along the azimuth has been reported in barn owls[7,8], ferrets[10], and Guinea pigs[11,12], but whether one exists in rodents has been debated. A study of auditory SC neurons employing anesthetized rats found only a weak correlation of the A–P positions and RF azimuths (r = 0.41, or r^2 = 0.17 in their report)[18]. Another study, which utilized anesthetized golden hamsters, found spatially restricted RFs only in the posterior SC[19]. We speculate that our use of awake-behaving animals is one reason that we found the spectral cue components of RF and a topographic map in mice, in contrast to the previous studies. Other improvements in our experimental design include: the use of the VAS system instead of a free-field speaker to deliver stimuli, which avoids acoustic interference of the sound with obstacles such as the recording rig; utilizing multichannel silicon probes to record from a large population of neurons in a single session increased the statistical power of our conclusions; thorough sampling of the auditory responses in two-dimensional spherical coordinates; RF parameter estimation through maximum likelihood model fitting using appropriate error analysis using quasi-Poisson statistics; and finally, separating the fast (< 20 ms) component of responses of the neurons. Each of these elements contributed to the accurate estimation of the topographic map properties in the present study.

Our results presented herein indicate that 23% of the auditory responsive SC neurons had a localized RF. This raises a question: what is the nature of the remaining 77% of auditory responsive neurons? Are they simply neurons whose RF was below our significance threshold, or do they have unique functions such as modulating the response properties? It is also known that visuoauditory integration neurons exist in the deep SC[20]; learning more about these neurons will give insights into mechanisms of sensory integration.

Roles of ITDs, ILDs and spectral cues in making a map of auditory space in the SC

We found that spectral cues are required for nasal RFs and ILDs are required for temporal RFs, while ITDs play no role in spatial computation in the mouse SC. Our finding that

spectral cues are necessary for generating spatially restricted RFs in the mouse SC should clarify their importance. Our work is consistent with that of Keating et al., who showed that ferrets can be trained to perform azimuthal sound localization with one ear plugged[5], and that of Huang and May, who showed that cats exhibit poor head orientation toward the sound when the mid-frequency spectral cues are eliminated from the stimulus[21]. However, spectral cues did not play a role in forming the RFs of neurons in the marmoset nucleus of the brachium of the inferior colliculus (nBIC), a major source of auditory input to the SC[4]. While we cannot directly compare these results with ours, future work in mice and other species will elucidate how, and under what conditions, spectral cues contribute to sound localization.

We found that most ILD-dependent neurons have RFs in the temporal area, close to where the ILD is maximized (Figs. 2o, 3l). These neurons are not tuned to a specific level difference but instead respond when level differences are greater than a given threshold (their response curve is a sigmoid-like monotonic function rather than a tuned, Gaussian-like non-monotonic response curve; Fig. 3o). This property is also found in the majority of neurons in the lateral superior olive (LSO) and the central nucleus of the inferior colliculus (ICC), the auditory nuclei that process ILDs[22]. Some non-monotonic neurons were found in the ICC of bats[22] and the SC of cats[23], but we found only one non-monotonic neuron out of the 79 ILD sensitive neurons. Therefore, when spectral cues are eliminated from the stimulus, most neurons change their RF azimuths to ~65° where the ILD is maximized (Figs. 2o, 3l, also see Fig. S1e for the ILD as a function of the location). Although these neurons may potentially encode different sound source azimuths by using different activation thresholds (as suggested[22]), they still make an RF in the same position in the absence of spectral cues. All these data indicate that SC neurons use ILDs in combination with spectral cues to differentiate their RF positions (Figs. 2k, 3k).

ITD elimination did not alter the RFs of the SC neurons. These results are consistent with previous reports, which concluded that ITDs do not contribute to the RF properties in ferrets[2] or marmosets[4], and the observation that the mouse has an underdeveloped medial superior olive[24], a nucleus that processes ITDs in other animals. This makes sense because the maximum ITD we measured for the mouse was only 29 μs (Fig. S1f). Therefore, it is unlikely that ITDs contribute to sound localization in mice, as generally predicted for small-headed mammals[25].

A potentially important difference between ILDs and spectral cues is that an appropriate interpretation of spectral cues requires knowledge of the original spectrum of the source sound. For a given stimulus, the ILD of a given source location is always the same (although it is a function of frequency). On the other hand, the spectra of the sound that arrived at the eardrums are influenced by both the spectra of the source sound and the HRTF, but the brain cannot tell them apart. (This is exactly why VAS stimulation works.) Therefore, if the spectrum of the source sound is an abnormal shape or restricted to a narrow frequency band, spectral cues will not be able to provide accurate information of the sound source location. It will be interesting to test the tuning properties and relative importance of the spectral cues and ILDs using different types of auditory stimuli, including tonal sound and naturalistic sound.

The nasal vs. temporal RF model explains the mechanisms used to compute auditory RFs and their topographic organization

Making a mathematical model of a biological system is useful for improving our understanding of the system and making testable predictions[26]. Our results lead to a model that incorporates the roles of spectral cues and ILDs, together with a baseline firing rate and a rectified linear function (Fig. 4a). This model recapitulates the formation of the auditory RFs, the heterogeneous use of spectral cues and ILDs across the SC, and the changes to the RFs when each cue is eliminated. Despite using a smaller number of parameters than used for the Kent distribution fit (4 compared to 7 per neuron, see Methods for details), this simple model describes ~82% of horizontal and ~53% of full-field RFs. However, this model does not provide a good description of elevation tuning, which is thought to rely on spectral cues[3]. A possible modification to the model would be the inclusion of terms that contain information to encode elevation, in particular a term with an RF at high elevation. The validity of the added elements can be verified by further experiments that eliminate specific components of the sound localization cues.

Conclusion

While ITDs and ILDs were thought to be the only important sources for sound localization in the horizontal plane, we discovered that spectral cues are responsible for nasal RFs, and are necessary for tuning RF azimuths and making a topographic map of sound source location in the mouse SC. It will be interesting to test whether spectral cues also play this unexpectedly important role in encoding azimuthal sound source location in other species, especially in larger mammals or birds that can utilize ITDs.

Materials and Methods

All procedures were performed in accordance with the University of California, Santa Cruz (UCSC) Institutional Animal Care and Use Committee.

Measurement of the head-related transfer function (HRTF)

To measure the HRTF of the mouse, a pair of Golay codes[27] (2^{16} = 65536 points for each) were played at a sampling rate of 500 kHz through a digital-to-analog converter (DAC; National Instruments (NI), PCIe-6341), an amplifier (Tucker Davis Technology (TDT), ED1), and an open-field electrostatic speaker (ES1) toward a decapitated mouse head located 25 cm away. The response was recorded with a microphone (Bruel & Kjaer, 4138-L-006, frequency response: 6.5 Hz to 140 kHz) coupled with the back of the ear canal of the mouse, amplified (Bruel & Kjaer, 1708), low-pass filtered (Thorlabs, EF502, 100 kHz) and digitized (NI, PCIe-6341) with a sampling rate of 500 kHz. The head-related impulse response (HRIR) was

reconstructed from the responses to the Golay codes[27]. The measurement was performed in an anechoic chamber (a cube with 60-cm-long sides) with egg-crate shaped polyurethane foam attached to the ceiling, walls, and the floor of the chamber as well as other surfaces such as the stage for the mouse head, the microphone and its cable, and the arm that holds the microphone.

Before measuring the HRTFs, the mouse was euthanized and then decapitated. The cochleae and the eardrums were removed from the ear canal and the microphone was coupled to the back of the ear canal using a ~1 cm long coupling tube. In order to keep the tissue from drying and ensure a continuous seal between the ear canal and the microphone, a small amount of mineral oil was applied on the tissue around the ear canal. To measure the HRTF as a function of the incident angle, the mouse head was mounted on a stage that was coupled to a stepper motor. The stage was rotated by the stepper motor and tilted by a hinge, without needing to detach the microphone (Fig. S1a). We measured the HRTF for grid points of 10 elevations (0° to 90° with 10° steps) and 101 azimuths (0° to 180° with 1.8° steps), 1010 points in total, in the upper right quadrant of the mouse head. Note that the polar axis of the coordinate for this measurement is the rostral direction (Fig. S6a). Before the microphone was detached, the HRTF of the setup with earphone (earphone HRTF, EHRTF) was also measured. This EHRTF was subtracted during the electrophysiological recording in order to avoid the additional frequency response due to the earphone set up. The same procedure was repeated for the left ear to achieve binaural HRTFs. Measuring the HRTFs from 3 mouse heads showed good reproducibility (the average difference was 5.3 ± 0.8 dB). We estimated the inconsistency of HRTFs between animals as an angular error (~±14°; see 'Estimation of additional systematic errors of the RF parameters' in 'Data Analysis') and incorporated this into our error measurements.

Virtual auditory space stimulation

Using the measured HRTFs, we developed a virtual auditory space (VAS) stimulation that creates stimulus sound with properties consistent with the sound coming from a specific source direction. This approach has been successfully used in humans[14], marmosets[4], Guinea pigs[11], and ferrets[2]. The stimulus sound is first filtered by a zero-phase inverse filter of the ES1 speaker and the EHRTF, to eliminate the non-flat frequency responses of the ES1 and the earphone setup. Next, the stimulus is filtered by a measured HRIR, including phase, to construct the sound property that is consistent with the incident angle for which the HRIR is measured. We only reproduced the frequency response in the range between 5 kHz and 80 kHz because the ES1 speaker did not produce a sufficient amplitude outside of this range. Within this range, the reconstructed sound reproduces the ITDs, ILDs, and spectral cues.

Sound is delivered to the ears through a DAC (NI, PCIe-6341), an amplifier (TDT, ED1), and closed-field electrostatic speakers (EC1) coupled with a small plastic horn. The tip of the horn was oriented toward the ear canal (~40° elevation, ~110° azimuth) and placed ~1 cm from it. With this angle, the EHRTF does not contain a strong notch and thus cancellation by an inverse filter was easy.

The full-field stimulus used grid points of 5 elevations (0° to 80° with 20° steps) and 17 azimuths (−144° to 144° with 18° steps), totaling 85 points in the two-dimensional directional field (Fig. S6a). We measured the HRTFs only in the upper right quadrant because of the limitation of the measurement stage. In order to construct the HRTFs in the upper left quadrant, we copied the transfer function of the opposite ear, flipping left and right. This is done after confirming that the left ear HRTF and right ear HRTF are similar in the horizontal plane (left-right symmetry of the ear shapes). The measured ITDs and ILDs for the upper right quadrant are plotted in Fig. S1e, f. The baseline stimulus pattern was 100-ms white noise with linear tapering windows in the first and last 5 ms. The stimulus intensity was 50 dB SPL. We generated a new pattern of white noise at every point of space and on every trial (the pattern is not 'frozen').

Because we used an open-air type earphone, some sound from the earphone is also detected by the contralateral ear. We measured this cross-talk amplitude to be ~−10 dB at 5 kHz and <−30 dB at 30 kHz and above, relative to the stimulated ear.

Cue elimination: To eliminate ITDs, we first measured the peak timings of the HRIR for each location in each ear and shifted the left HRIR to eliminate the timing difference. To eliminate ILDs, we calculated the average amplitudes between 5–80 kHz in left and right HRTFs and adjusted their amplitudes to be the average of the left and right HRTFs so that they have equal overall sound levels across this range of frequency. To eliminate spectral cues, we replaced the HRIR with a single sample impulse to flatten the spectrum, while the overall sound level and timing were kept consistent with the original sound. To create monaural stimuli, we simply sent no signal to the ipsilateral (left) speaker. To test whether the neurons are tuned to larger ILD values, we used the same 100-ms white noise bursts with a flat spectrum but changed the ILD randomly between ±40 dB with 2 dB increments. The stimulus for each ILD value was repeated 30 times.

Animal Preparation for electrophysiology

We used 2–5 month-old CBA/CaJ (The Jackson Laboratory, 000654) mice of each sex. One day before the recording, we implanted a custom-made titanium head plate on a mouse's skull, which allowed us to fix the mouse's head to the recording rig without touching the ears. On the day of the recording, the mouse was anesthetized with isoflurane (3% induction, 1.5–2% maintenance; in 100% oxygen) and a craniotomy was made (~1.5–2 mm diameter) in the left hemisphere above the SC (0.6 mm lateral from the midline, on the lambdoid suture). The mouse was given >1 hour to recover from the anesthesia before recording. The incision was covered with 2% low-melting-point agarose in saline and a layer of mineral oil on top of it to keep the brain from drying. A 256-channel silicon probe (provided by Prof. Masmanidis[28]) was inserted through the cortex into the SC with its four shanks aligned along the A–P axis (Fig. 1b; the electrodes were facing toward the lateral direction). When the shank enters the superficial SC, the positions of multiunit visual RFs on each shank are recorded. We then lower the probe until the visual responses disappear. As a consequence, the top of the active area of the most superficial shank is located at ~300 μm from the surface of the SC. Before fixing the probe at the final location, the probe was overshot by ~120 μm and rewound in order to reduce probe

migration during the recording. Recordings were started 20–30 min after inserting the probe. The mice were euthanized after the recording session.

During the recording, the mouse was allowed to run freely on a cylindrical treadmill made of polystyrene foam (Fig. 1a). The surface of the cylinder was covered with self-adherent wrap (CVS pharmacy) to reduce locomotion noise. The movement of the cylinder was recorded by a shaft encoder (US Digital, H5-100-NE-S).

In 5 of 10 experiments in the “Mice have spatially localized auditory RFs that are topographically organized” section, we inserted the probe 3 times per mouse in order to span the recordings in different mediolateral (M–L) positions to measure the elevation tuning. Each penetration was separated by 400 μm along the M–L axis. The order of penetration in the M–L axis was randomized for each experiment to avoid a correlation between a degradation in recording quality with the probe location. All the other mice had only one probe insertion.

Data analysis

Blind analysis

We performed a blind analysis of our data, which is effective for reducing a false-positive reporting of results[29,30]. First, we looked for specific features using half of the neurons randomly subsampled from each experiment (exploratory dataset). After deciding which parameters to look at, we performed the same analysis on the hidden half of the data (blinded dataset) to confirm if the results agreed with the exploratory dataset. All of the results reported in the present study passed a significance test both in the exploratory dataset and the blinded dataset unless noted otherwise.

Estimation of the anteroposterior locations of the neurons

We estimated the relative position of the silicon probes in multiple experiments using the visual RF positions measured in the superficial SC. When we inserted a silicon probe, we measured the positions of the visual RFs on each shank in the superficial SC using multi-unit activity. We extrapolated the visual RFs to find an A–P position where the visual RF azimuth was 0°, and defined this point as the zero of the A–P position. To estimate the A–P position of the auditory neurons we assumed: (1) the most superficial electrode of the silicon probe was at 300 μm in depth; and (2) the insertion angle of the probe into the SC was 25°. (1) is justified by consistently stopping penetration at the position where visual responses predominantly disappear from all of the electrodes; and (2) is justified by the post-recording observations of the insertion angle (Fig. 1c). The position of a neuron relative to the probe was determined by a two-dimensional Gaussian fit to its spike amplitude across multiple electrodes[31]. We only analyzed neurons with a positive A–P position in order not to include neurons outside the SC.

Significance test for the auditory responses

We used quasi-Poisson statistics for significance tests of the auditory responses of individual neurons[32]. Simple Poisson statistics was not sufficient because the post-stimulus firing rate typically had a larger trial-by-trial variance than that expected from Poisson statistics (overdispersion) due to factors such as bursting of the neural activity and/or locomotion/movement of the animal. These additional fluctuations can cause increased false positives. Therefore, we first estimated an overdispersion parameter by calculating the standard deviation of the spike counts and evaluated the significance using quasi-Poisson statistics.

Function fit to estimate the azimuth and elevation of the RFs of the neurons

We used a maximum likelihood fit of the Kent distribution[15] to estimate the azimuth, the elevation, and the radius of an RF. The Kent distribution is a spatially localized distribution in a two-dimensional directional space.

The equation of the Kent distribution is given by:

$$f(\vec{x}) = \frac{1}{\exp(\kappa)} \exp\{\kappa\vec{\gamma_1} \cdot \vec{x} + \beta[(\vec{\gamma_2} \cdot \vec{x})^2 - (\vec{\gamma_3} \cdot \vec{x})^2]\}$$

where $\vec{x}$ is a three-dimensional unit vector; κ is a concentration parameter that represents the size of the RF ($\kappa > 0$); β is an ellipticity parameter ($0 < 2\beta < \kappa$); the vector $\vec{\gamma_1}$ is the mean direction of the RF; vectors $\vec{\gamma_2}$ and $\vec{\gamma_3}$ are the major and minor axes of the Kent distribution. These three unit vectors are orthogonal to each other. This function is normalized to 1 when $\vec{x} = \vec{\gamma_1}$. At the limit of large κ, this distribution becomes asymptotically close to a two-dimensional Gaussian. At each point of the directional field, the likelihood value was calculated based on quasi-Poisson statistics[32].

We used the front-Z coordinate system (Fig. S6a) for the HRTF measurement and the stimulus presentation but used the top-Z coordinate system (Fig. S6b) for the Kent distribution fitting. We had to use the front-Z coordinate system for the HRFT measurement because of the restriction of our measurement stage (Fig. S1a). However, the front-Z coordinate system has a discontinuity of the azimuth across the midline that gives a problem in fitting and interpretation of the data near the midline. We avoided this issue by switching to the top-Z coordinate system. Because the Kent distribution in vector format is independent of the coordinate system, likelihood values and parameters in the vector format are not affected by this change when the fit is successful. To avoid an area where two coordinate systems are largely different, we only used neurons with their elevation smaller than 30° for azimuthal topography (Fig. 1h). To achieve stable fits, we restricted β to satisfy $4\beta < \kappa$, set the azimuthal range to be from −144° to +144°, and set the elevation range to be from 0° to 90°.

We used the following equation to define the radius ρ from the concentration parameter κ.

$$\rho = \cos^{-1}\left(1 - \frac{1}{2\kappa}\right)$$

With this definition, the value of the distribution becomes $\exp(-\frac{1}{2})$ of the peak value when $\left|\vec{x} \cdot \vec{\gamma_1}\right| = \cos(\rho)$ and β is negligible, a behavior similar to a two-dimensional Gaussian.

Estimation of additional systematic errors of the RF parameters

We estimated additional systematic errors that may affect the RF parameters. First, the mouse was alert and freely running on a cylindrical treadmill during the neural recording session. Locomotion modulates the functional properties of the visually responsive neurons[31,33,34] and the auditory cortical neurons[35,36]. To determine the effects of locomotion on the auditory SC neurons, we separated the session into segments when the mouse was running (speed >1 cm/s) and when it was stationary. The mouse was running, on average, 30% of the time (Fig S5a). During locomotion, the spontaneous activity increased and the auditory responses decreased both in the fast and slow timescales (Fig. S5b–g). However, it did not change the individual RF structures or the map of auditory space (Fig. S5h, i). Therefore, we did not consider this as a source of a systematic error. Second, the mouse was moving its eyes during the experiment. The superficial SC receives direct retinal projections and eye movement modulates the map of auditory space to keep the visual and auditory maps aligned[37]. However, the range of the mouse eye movement is small. The standard deviation (SD) is ~2–3° and the saccade amplitude is < 10° along the horizontal axis and the movement along the vertical axis is even smaller[33,38]. Therefore, it is unlikely that this is a major source of systematic errors. Based on the standard deviation value, we take ±3° to be the estimated systematic error. Third, the HRTFs can be modulated by movement of the mouse pinnae. To estimate the effect of the pinna movement on the HRTF, we measured the amplitude of the ear movement using an infrared video camera (Basler acA640-120um) while the mouse listened to the auditory stimuli. We quantified the pinna angle in 100 randomly selected images of the video recording. Overall, the mouse moved the pinna by 13° (SD). Fourth, our VAS stimuli might not be perfect because we used a single set of HRTFs for all the mice instead of measuring them individually. In order to quantify this effect, we measured HRTFs from three mice and compared the differences (Fig. S1b). The overall difference (RMS) was 5.3 ± 0.8 dB, and this corresponds to an angular shift of ±14°.

To estimate the topographic map parameters, taking into account both the systematic and statistical errors, the systematic errors noted above were added in quadrature to give ±19°. This is larger than the average statistical errors associated with the individual neurons (±5.8°). However, a linear fit to the topographic map along the azimuthal axis (Fig. 1h) resulted in a large χ^2 per degree of freedom (24.1), which is in the regime of either underfitting or underestimation of the error. In order to estimate the slope and offset parameters, we added, in quadrature, the estimated systematic error of ± 19° to each data point. The χ^2 per degree of freedom, based on a linear fit, is now 0.69, indicating that the systematic errors are a reasonable addition to the statistical errors.

Comparing the effect of the elimination of sound localization cues on the RFs of SC neurons

We used cosine similarity to compare the similarity of the auditory RFs before and after each cue elimination. We considered the auditory responses at 85 virtual source locations as an

85-dimensional vector and normalized it so that it has a unit length and zero mean. Then we took an inner product of the responses to the original stimulus and the additional stimuli (control, no ITD, no ILD, no spectral cue) to calculate the cosine similarity indices (SI). The statistical errors of the spike count at each virtual source location were propagated to the final error of the SIs. The SIs of the cue-elimination experiments are further normalized by dividing by the SI of the control stimuli (NSI; normalized SI).

Modeling receptive fields of mouse SC neurons

We developed a model of the RF structure based on our data. The model equation is the following:

$$F(\vec{x}) = \max\left(b + w_1 F_{spec}(\vec{x};\kappa) + w_2 F_{ILD}(\vec{x}),\ 0\right)$$

where F_{spec} is a Kent distribution with $\beta = 0$ and γ_1 pointing directly to the front, F_{ILD} is the measured ILD at each location averaged over the range of 5–80 kHz in dB (Fig. S1e), and normalized to have a range of −1 to 1. The baseline firing rate (b), weight for F_{spec} (w_1), weight for F_{ILD} (w_2), and a parameter of the Kent distribution that is related to the spectral RF size (κ) are the model parameters (Fig. 4a). Max function is a function that takes the larger of the two arguments, acting as a rectified linear function as shown in Fig 4a. This function was fit to the firing rates in the latency period between 5–20 ms using the same maximum likelihood method with quasi-Poisson statistics that is described above. The goodness-of-fit significance was estimated using the negative log-likelihood multiplied by two based on Wilks' theorem[39].

References

1. Rayleigh, Lord. XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **13**, 214–232 (1907).
2. Campbell, R. A. A., Doubell, T. P., Nodal, F. R., Schnupp, J. W. H. & King, A. J. Interaural timing cues do not contribute to the map of space in the ferret superior colliculus: a virtual acoustic space study. *J. Neurophysiol.* **95**, 242–254 (2006).
3. Grothe, B., Pecka, M. & McAlpine, D. Mechanisms of sound localization in mammals. *Physiol. Rev.* **90**, 983–1012 (2010).
4. Slee, S. J. & Young, E. D. Linear processing of interaural level difference underlies spatial tuning in the nucleus of the brachium of the inferior colliculus. *J. Neurosci.* **33**, 3891–3904 (2013).
5. Keating, P., Dahmen, J. C. & King, A. J. Context-specific reweighting of auditory spatial cues following altered experience during development. *Curr. Biol.* **23**, 1291–1299 (2013).
6. Slattery, W. H., 3rd & Middlebrooks, J. C. Monaural sound localization: acute versus chronic unilateral impairment. *Hear. Res.* **75**, 38–46 (1994).
7. Knudsen, E. I. Auditory and visual maps of space in the optic tectum of the owl. *J. Neurosci.* **2**, 1177–1194 (1982).
8. Knudsen, E. I. Auditory properties of space-tuned units in owl's optic tectum. *J.*

Neurophysiol. **52**, 709–723 (1984).
9. Middlebrooks, J. C. & Knudsen, E. I. A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.* **4**, 2621–2634 (1984).
10. King, A. J. & Hutchings, M. E. Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space. *J. Neurophysiol.* **57**, 596–624 (1987).
11. Sterbing, S. J., Hartung, K. & Hoffmann, K.-P. Representation of sound source direction in the superior colliculus of the guinea pig in a virtual auditory environment. *Exp. Brain Res.* **142**, 570–577 (2002).
12. King, A. J. & Palmer, A. R. Cells responsive to free-field auditory stimuli in guinea-pig superior colliculus: distribution and response properties. *J. Physiol.* **342**, 361–381 (1983).
13. Cang, J. & Feldheim, D. A. Developmental mechanisms of topographic map formation and alignment. *Annu. Rev. Neurosci.* **36**, 51–77 (2013).
14. Wightman, F. L. & Kistler, D. J. Headphone simulation of free-field listening. I: Stimulus synthesis. *J. Acoust. Soc. Am.* **85**, 858–867 (1989).
15. Kent, J. T. The Fisher-Bingham Distribution on the Sphere. *J. R. Stat. Soc. Series B Stat. Methodol.* **44**, 71–80 (1982).
16. Lauer, A. M., Engel, J. H. & Schrode, K. Rodent Sound Localization and Spatial Hearing. in *Rodent Bioacoustics* (eds. Dent, M. L., Fay, R. R. & Popper, A. N.) 107–130 (Springer International Publishing, 2018).
17. Lauer, A. M., Slee, S. J. & May, B. J. Acoustic basis of directional acuity in laboratory mice. *J. Assoc. Res. Otolaryngol.* **12**, 633–645 (2011).
18. Gaese, B. H. & Johnen, A. Coding for auditory space in the superior colliculus of the rat. *Eur. J. Neurosci.* **12**, 1739–1752 (2000).
19. Chalupa, L. M. & Rhoades, R. W. Responses of visual, somatosensory, and auditory neurones in the golden hamster's superior colliculus. *J. Physiol.* **270**, 595–626 (1977).
20. Stein, B. E., Stanford, T. R. & Rowland, B. A. Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* **15**, 520–535 (2014).
21. Huang, A. Y. & May, B. J. Sound orientation behavior in cats. II. Mid-frequency spectral cues for sound localization. *J. Acoust. Soc. Am.* (1996).
22. Park, T. J., Klug, A., Holinstat, M. & Grothe, B. Interaural level difference processing in the lateral superior olive and the inferior colliculus. *J. Neurophysiol.* **92**, 289–301 (2004).
23. Wise, L. Z. & Irvine, D. R. Topographic organization of interaural intensity difference sensitivity in deep layers of cat superior colliculus: implications for auditory spatial representation. *J. Neurophysiol.* **54**, 185–211 (1985).
24. Irving, R. & Harrison, J. M. The superior olivary complex and audition: a comparative study. *J. Comp. Neurol.* **130**, 77–86 (1967).
25. Grothe, B. & Park, T. J. Sensitivity to interaural time differences in the medial superior olive of a small mammal, the Mexican free-tailed bat. *J. Neurosci.* **18**, 6608–6622 (1998).
26. Goodhill, G. J. Theoretical Models of Neural Development. *iScience* **8**, 183–199 (2018).
27. Zhou, B., Green, D. M. & Middlebrooks, J. C. Characterization of external ear impulse responses using Golay codes. *J. Acoust. Soc. Am.* **92**, 1169–1171 (1992).
28. Shobe, J. L., Claar, L. D., Parhami, S., Bakhurin, K. I. & Masmanidis, S. C. Brain activity

mapping at multiple scales with silicon microprobes containing 1,024 electrodes. *Journal of Neurophysiology* **114**, 2043–2052 (2015).
29. MacCoun, R. & Perlmutter, S. Blind analysis: Hide results to seek the truth. *Nature* **526**, 187–189 (2015).
30. Klein, J. R. & Roodman, A. BLIND ANALYSIS IN NUCLEAR AND PARTICLE PHYSICS. *Annu. Rev. Nucl. Part. Sci.* **55**, 141–163 (2005).
31. Ito, S., Feldheim, D. A. & Litke, A. M. Segregation of Visual Response Properties in the Mouse Superior Colliculus and Their Modulation during Locomotion. *J. Neurosci.* **37**, 8428–8443 (2017).
32. Ver Hoef, J. M. & Boveng, P. L. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology* **88**, 2766–2772 (2007).
33. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
34. Fu, Y., Kaneko, M., Tang, Y., Alvarez-Buylla, A. & Stryker, M. P. A cortical disinhibitory circuit for enhancing adult plasticity. *Elife* **4**, e05558 (2015).
35. Schneider, D. M., Nelson, A. & Mooney, R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–194 (2014).
36. Schneider, D. M., Sundararajan, J. & Mooney, R. A cortical filter that learns to suppress the acoustic consequences of movement. *Nature* **561**, 391–395 (2018).
37. Jay, M. F. & Sparks, D. L. Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J. Neurophysiol.* **57**, 35–55 (1987).
38. Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L. & Reid, R. C. Functional specialization of mouse higher visual cortical areas. *Neuron* **72**, 1025–1039 (2011).
39. Wilks, S. S. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Ann. Math. Stat.* **9**, 60–62 (1938).
40. Einevoll, G. T., Kayser, C., Logothetis, N. K. & Panzeri, S. Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat. Rev. Neurosci.* **14**, 770–785 (2013).

Acknowledgements

This work was supported by Brain Research Seed Funding provided by UCSC, the National Institutes of Health (Grant NEI R21EY026758 to D.A.F. and A.M.L.), a Whitehall foundation grant to D.A.F., and a donation from John Chen to A.M.L.. We thank Sotiris Masmanidis for providing us with the silicon probes; Ben Abrams, Director of the UCSC Life Science Microscopy Center, for training and assistance with imaging; J. Gerard G. Borst, Jena Yamada, Sadaf Abed, and Youngmin Yu for helpful comments on the manuscript.

Author Information

Affiliations

Santa Cruz Institute for Particle Physics, University of California, CA, Santa Cruz, USA
Shinya Ito & Alan M. Litke
Department of Molecular, Cell and Developmental Biology, University of California, CA, Santa Cruz, USA
Yufei Si & David A. Feldheim

Contributions

S.I., D.A.F. and A.M.L. conceived the project. S.I. designed the experiments. S.I. and Y.S. acquired the data. S.I. and Y.S. analyzed the data. S.I., D.A.F. and A.M.L interpreted the results and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Corresponding author

Correspondence to Shinya Ito and Alan M. Litke.