Supporting Information

Atencio et al. 10.1073/pnas.0908383106

SI Text

Electrophysiology. Electrophysiological methods and stimulus design were similar to those described (1, 2). Young adult cats were given an initial dose of ketamine (22 mg/kg) and acepromazine (0.11 mg/kg), and anesthesia was maintained with pentobarbital sodium (Nembutal, 15–30 mg/kg) during the surgical procedure. The animal's temperature was maintained with a thermostatic heating pad. Bupivicaine was applied to incisions and pressure points. Surgery consisted of a tracheotomy, reflection of the soft tissues of the scalp, craniotomy over AI, and durotomy. After surgery, the animal was maintained in an areflexive state with a continuous infusion of ketamine/diazepam (2–5 mg/kg per hr ketamine, 0.05–0.2 mg/kg per hr diazepam in lactated Ringer's solution). All procedures were in strict accordance with the policies of the University of California, San Francisco Committee for Animal Research.

With the animal in a sound-shielded anechoic chamber (IAC), stimuli were delivered via a closed speaker system (diaphragms from Stax). Multichannel silicon recording probes (kindly provided by the University of Michigan Center for Neural Communication Technology) were used to make simultaneous extracellular recordings. The probes had 16 linearly spaced recording channels, with each channel separated by 0.15 mm. The impedance of each channel was 2–3 MΩ. Probes were carefully positioned orthogonally to the cortical surface and lowered to depths between 2.3 and 2.4 mm by using a microdrive (David Kopf Instruments).

Recording Procedure. All recording locations were in AI, as verified through initial multiunit mapping and determined by the layout of the tonotopic gradient and bandwidth modules on the ectosylvian gyrus (3). For each animal, a digital photo was acquired that contained the suprasylvian, the posterior ectosylvian, and the anterior ectosylvian sulci (Fig. S1). The image was imported into Canvas software (ACD Systems). We then drew a line on the image, which connected the tips of the posterior and the anterior ectosylvian sulci. In cat AI, this line indicates the ventral boundary of the central narrowly tuned region (4, 5); ≈2 mm dorsal to this line is the dorsal boundary of the central narrowly tuned region. Recordings were obtained in this central region. We then made another line on the image, which bisected, and was orthogonal to, the line that connected the tips of the sulci. Penetrations were made along this ventral–dorsal line in AI. Before each penetration, we adjusted the multichannel array by using a micromanipulator (Narshige). The array was visualized along the dorsal–ventral and then the anterior–posterior axes of AI, to ensure that the insertion was orthogonal to the cortical surface. After insertion, each penetration was marked on the digital image. Before recording, at least 20 min passed. The first stimulus presented was a tuning curve stimulus that lasted ≈17 min. Thus, the receptive field data presented in this study were collected ≈40 min or more after electrode insertion, which allowed cortical "dimpling" effects to be minimized.

This careful procedure helped ensure that the linear recording array was orthogonal to the cortical surface and spanned all cortical layers. Anatomical variation effects were minimized with this approach, because the penetrations were geographically removed from each sulcus, and they were on the center of the ectosylvian gyrus, which is relatively flat in cat AI (Fig. S1). Operationally, we refer to this recording approach as "columnar," which implies that the activity of recorded neurons represents processing that spans the full vertical thickness of the cortical laminae, but may include more interactions than represented by the extent of anatomical microcolumns, and less than the extent of functional modules (6). Recording depths were anatomically confirmed in selected cases. In each case, neurons recorded at depths between 0.6 and 1.1 mm had the shortest receptive field minimum latencies, and the depths of the neurons (0.6 to 1.1 mm) coincided with the anatomical location of granular layers 3B/4. After this confirmation, we predominantly used depth recordings to localize the position of single units.

To further confirm the orthogonal nature of the penetrations, we examined the depth distribution of characteristic frequency (CF) and latency. How these parameters varied with depth provided further evidence for the orthogonality of the penetrations. First, the frequency preference of the neurons (with the occasional exception for the deepest locations) was constant. Across all penetrations in this study, the average CF disparity was 0.1 ± 0.1 octaves (mean ± SD) in a penetration (7). Second, the latency profile matched the accepted pattern, where the shortest values were found around thalamic input layers 3B/4.

Neural traces were bandpass-filtered between 600 and 6,000 Hz and recorded to disk with a Neuralynx Cheetah A/D system at sampling rates between 18 and 27 kHz. The traces were sorted off-line with a Bayesian spike-sorting algorithm (8). Each probe penetration yielded 8–16 active channels, with ≈1–2 single units per channel. Stimulus-driven neural activity was recorded for ≈75 min at each location.

Stimulus. For any recording position, neurons were probed with pure tones, then one or two presentations of a 15 or 20-min dynamic moving ripple stimulus. Using the 15-min stimulus allowed us to obtain multiple responses to the ripple stimulus, which were used for later cross-correlation analysis (20). The dynamic ripple stimulus was a temporally varying broadband sound (500–20,000 or 40,000 Hz) composed of ≈50 sinusoidal carriers per octave, each with randomized phase (9). The 20-kHz ripple was used at sites with CFs <12 kHz. The 40 kHz ripple stimulus was used at higher CF sites. The magnitude of the logarithmically spaced carriers was modulated by the spectrotemporal envelope, which was defined by spectral (0–4 cycles/octave) and temporal (−40 to 40 cycles/s) modulation parameters. Maximum modulation depth of the spectrotemporal envelope was 40 dB. Mean intensity was set at 70 or 80 dB SPL, which allowed us to set the intensity ≈40 dB above the minimum threshold at a multicontact recording site.

Analysis. Data analysis was carried out in MATLAB (Mathworks). We used two techniques to derive auditory STRFs. In the first, we used the reverse correlation method to derive the spike-triggered average of the spectrotemporal stimulus envelope immediately preceding a spike (STA) (9–13).

The second technique we used to estimate STRFs was the method of MIDs, in which we followed previously reported methodologies (14, 15). To find STRFs, we searched through the stimulus space for those dimensions that maximized the MI between the stimulus and the spiking response. The first MID (MID1) is the STRF that accounts for the most MI between the stimulus and the response. The second MID (MID2) was then found as the STRF that, together with the first MID, further maximized the information. The MI between projections onto an individual STRF, v, and single spikes was computed according to

$$I(v) = \int dx P_v(x|spike) \log_2\left[\frac{P_v(x|spike)}{P_v(x)}\right],$$

where x represents projections onto the STRF, v; subscript v denotes the fact that the probability distributions change depending on the STRF v. The STRF v was either MID1 or MID2. The 1D nonlinearity was calculated via

$$P(spike|x) = P(spike)\frac{P(x|spike)}{P(x)}.$$

The 1D nonlinearity describes the spiking probability of a neuron as a function of how well the stimulus matches the receptive field. If the stimulus is similar to the receptive field ($x > 0$), then the probability of a spike will increase. If the receptive field is not similar to the stimulus ($x \approx 0$), then the probability of spiking decreases.

The MI between single spikes and both MIDs was calculated as

$$I(MID1, MID1) = \iint dx_1 dx_2 P(x_1, x_2|spike) \log_2$$
$$\left[\frac{P(x_1, x_2|spike)}{P(x_1, x_2)}\right],$$

where x_1 and x_2 represent the projections of the stimulus onto the first and second MIDs, respectively. The 2D nonlinearity was calculated via

$$P(spike|x_1, x_2) = P(spike)\frac{P(x_1, x_2|spike)}{P(x_1, x_2)}.$$

All estimates of relevant stimulus dimensions (STA, MID1, MID2) were computed by separating the data four different ways into training and test datasets. Each training set consisted of ¾ of the data, and each test set consisted of ¼ of the data (16). This resulted in four different estimates for each STRF. Information values were calculated by using different fractions of the test dataset for each neuron. To accomplish this, the information values were calculated over the first 80%, 90%, 92.5%, 95%, 97.5%, and 100% of the test dataset. The information calculated from these data fractions was plotted against the inverse of the data fraction percentage (1/80, 1/90, etc.). We extrapolated the information values to infinite dataset size by fitting a line to the plot and taking the ordinate intersect as the information value for unlimited data size.

The separability of the STA, the MIDs, and the 2D nonlinearity was determined by performing singular value decomposition (17). The separability index was defined as

$$SPI = \sigma_1^2 \Big/ \sum_i \sigma_i^2,$$

where σ_1 is the largest singular value. The SPI, ranges between 0 and 1, with 1 indicating that the function can be described as a product of two 1D functions. For STAs and MIDs, the SPI quantifies how well frequency and time can be dissociated in the receptive field. For a neuron with a symmetric or circular excitatory region, frequency and time are relatively independent in the receptive field, and the SPI would be near 1. For a neuron with more diagonally oriented, or asymmetric, excitatory or inhibitory regions, the SPI would be nearer to 0.

We computed the phase-locking index (PLI) for each neuron by using the relation $PLI = (\max(STA) - \min(STA))/(r\sqrt{8})$, where $\max(STA)$ and $\min(STA)$ are the maximum and minimum

values in the STRF, and r is the average firing rate (9). The PLI ranges from 0 (not phase locked) to 1 (precisely phase locked).

To determine the stimulus selectivity of each neuron we calculated a feature selectivity index (FSI) using previously published methodologies (9, 18). For each spike generated by the neuron, the ripple envelope that preceded the spike was captured and correlated with the neuron's STA. First, we defined a similarity index, SI, as

$$SI = \frac{\sum_i \sum_j stim(i,j)STA(i,j)}{\sqrt{\sum_i \sum_j stim(i,j)stim(i,j)}\sqrt{\sum_i \sum_j STA(i,j)STA(i,j)}},$$

where $stim$ and STA are matrices that represent the stimulus segment preceding a spike, and the STA of the neuron, respectively, and i and j range over the number of rows and columns in the STA. The SI ranges between $+1$ and -1 and is a measure of the spectrotemporal correlation between the stimulus and the STA. The SI is equivalent to the Pearson correlation coefficient.

We calculated a similarity index for each action potential, forming a SI probability distribution, $p(SI)$, of the driven activity. Using a spike train of similar length but from random spikes (1, 18), we calculated SIs from the neuron's STA and formed a probability distribution, $p_{rand}(SI)$, for a random selection of stimulus segments. For each SI probability distribution we calculated the cumulative distribution function according to

$$P(SI) = \int_{-1}^{SI} p(x)dx.$$

The difference between the random and driven spike trains was quantified by obtaining the areas, A and A_{rand}, under each cumulative distribution function. We then calculated the FSI as

$$FSI = \frac{A_{rand} - A}{A_{rand}}.$$

FSI values vary between 0 and 1, where 0 corresponds to similar distributions for $P_{rand}(SI)$, and $P(SI)$, i.e., a neuron that responds indiscriminately to stimulus segments, and 1 corresponds to a neuron that responds to a very restricted range of stimulus features. The FSI describes the stimulus integration properties of a neuron. It quantifies the variability or range of stimulus configurations that the neuron responds to. Higher FSI values imply that the neuron is more stimulus selective, responding only to tight matches between the stimulus and the receptive field, e.g., because of a higher threshold for spiking. Lower FSI values imply that the neuron integrates and responds to a wider range of stimuli, and thus the variability in, and range of, the stimulus set is greater for low FSI neurons than for high FSI neurons.

The first MID contribution was defined as

$$100\frac{I(MID1)}{I(MID1, MID2)}.$$

The synergy between the two MIDs was defined as

$$100\frac{I(MID1, MID2)}{I(MID1) + I(MID2)}.$$

For both measures we used the information values from the extrapolation procedure.

The shape of each 1D nonlinearity was characterized by an asymmetry index (ASI). The ASI is defined as $(R-L)/(R+L)$, where R and L are the sums of nonlinearity values that correspond to projection values > 0 or < 0, respectively. Values near

0 indicate a symmetric nonlinearity, implying that the neuron responds regardless of the stimulus phase or polarity. ASIs near $+1$ or -1 indicate neurons that have an increased probability of spiking when the stimulus is either positively or negatively correlated with the filter, respectively. Neurons that only fire when there is a positive correlation between stimulus and receptive field (i.e., high stimulus energy falls onto excitatory regions and low energy on inhibitory regions) have a highly asymmetric nonlinearity or firing probability function. Those neurons do not fire when the stimulus is negatively correlated with the STRF, i.e., with low energy falling on excitatory regions and high energy falling on inhibitory regions. Neurons that respond when stimulus–STRF correlations are either positive or negative have a more symmetric nonlinearity (values near 0). These neurons tend to have a lower FSI because they respond independently of the sign of the envelope phase.

1. Miller LM, Escabí MA, Read HL, Schreiner CE (2002) Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J Neurophysiol* 87:516–527.
2. Miller LM, Schreiner CE (2000) Stimulus-based state control in the thalamocortical system. *J Neurosci* 20:7011–7016.
3. Imaizumi K, Schreiner CE (2007) Spatial interaction between spectral integration and frequency gradient in primary auditory cortex. *J Neurophysiol* 98:2933–2942.
4. Read HL, Winer JA, Schreiner CE (2001) Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc Natl Acad Sci USA* 98:8042–8047.
5. Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: Distribution of integrated excitation. *J Neurophysiol* 64:1442–1459.
6. Linden JF, Schreiner CE (2003) Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. *Cereb Cortex* 13:83–89.
7. Abeles M, Goldstein MH, Jr (1970) Functional architecture in cat primary auditory cortex: Columnar organization and organization according to depth. *J Neurophysiol* 33:172–187.
8. Lewicki MS (1994) Bayesian modeling and classification of neural signals. *Neural Comput* 6:1005–1030.
9. Escabí MA, Schreiner CE (2002) Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *J Neurosci* 22:4114–4131.
10. Aertsen A, Johannesma PI (1980) Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli. *Biol Cybern* 38:223–234.
11. deCharms RC, Blake DT, Merzenich MM (1998) Optimizing sound features for cortical neurons. *Science* 280:1439–1443.
12. Klein DJ, Depireux DA, Simon JZ, Shamma SA (2000) Robust spectrotemporal reverse correlation for the auditory system: Optimizing stimulus design. *J Comput Neurosci* 9:85–111.
13. Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20:2315–2331.
14. Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Comput* 16:223–250.
15. Sharpee TO, et al. (2006) Adaptive filtering enhances information transmission in visual cortex. *Nature* 439:936–942.
16. David SV, Gallant JL (2005) Predicting neuronal responses during natural vision. *Network* 16:239–260.
17. Depireux DA, Simon JZ, Klein DJ, Shamma SA (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 85:1220–1234.
18. Miller LM, Escabí MA, Schreiner CE (2001) Feature selectivity and interneuronal cooperation in the thalamocortical system. *J Neurosci* 21:8136–8144.
19. Harman HH (1976) *Modern Factor Analysis* (Univ Chicago Press, Chicago), 3rd Ed.
20. Atencio CA, Schreiner CE (2007) Laminar distribution of temporal interactions in the primary auditory cortex of the cat. *Soc Neurosci Abstr* 278.5/EE4.
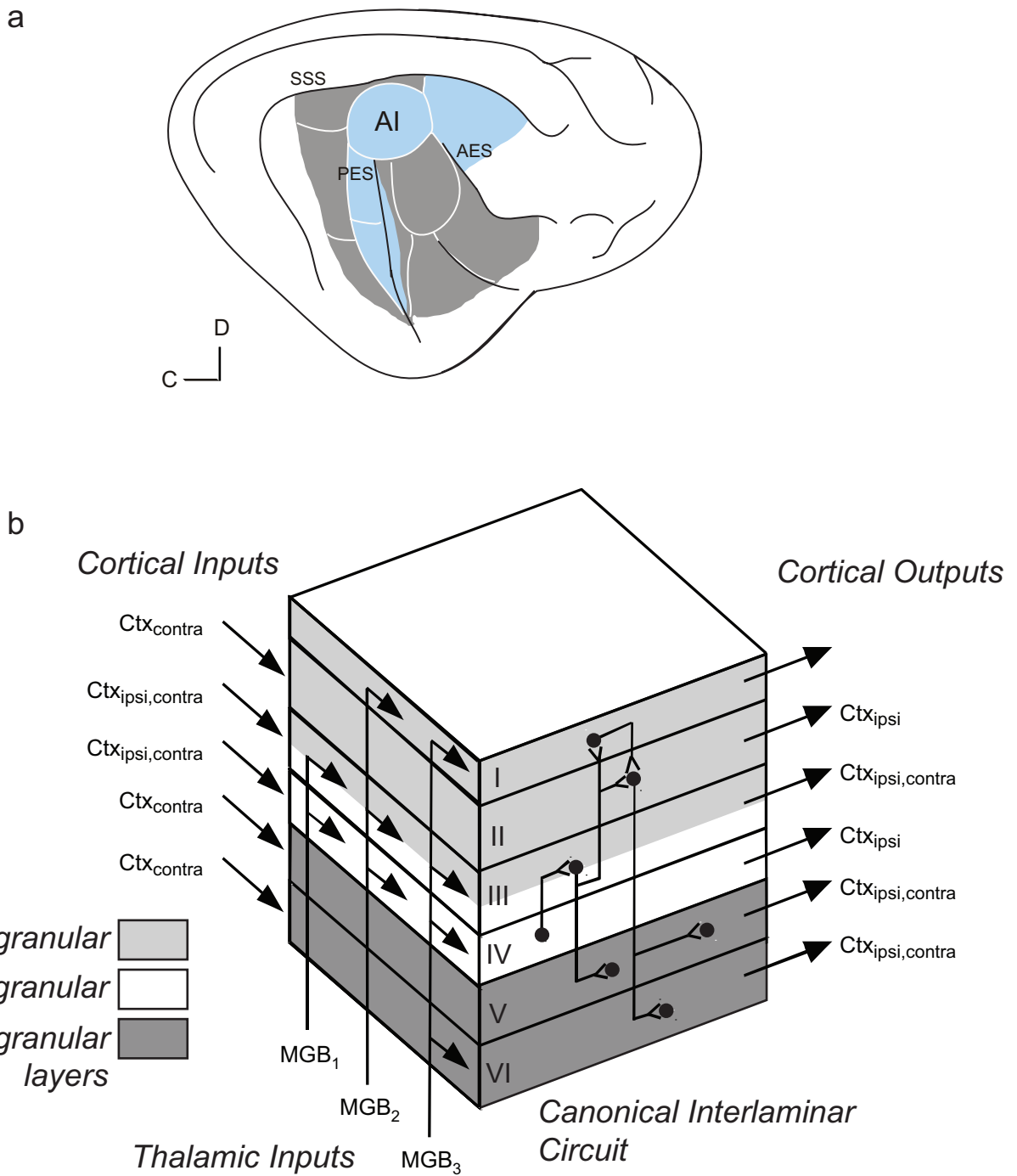
Fig. S1. Auditory cortical areas. (*a*) Auditory responsive areas in the cat brain. Gray signifies areas that respond to auditory stimuli, but do not have tonotopic organization. Blue denotes fields that have a tonotopic organization. White lines depict field boundaries. Black lines depict sulci. Three main sulci provide geographic boundaries for AI: SSS (suprasylvian sulcus); PES (posterior ectosylvian sulcus); and AES (anterior ectosylvian sulcus). AI lies on the gyrus between PES and AES. D: dorsal; C: caudal. Modified from ref. 1. (*b*) Schematic of columnar inputs and outputs to/from AI. A representative section of cortex, spanning the full cortical thickness, will receive cortical (Ctx) inputs from ipsilateral (ipsi) and contralateral (contra) hemispheres. Three types of thalamic inputs segregate according to layer, with MGB_1 representing lemniscal input to granular layers, and MGB_2 and MGB_3 coming from nonlemniscal auditory thalamus (2). Outputs from AI proceed to other regions of AI, and other cortical fields, in both the ipsilateral and contralateral hemispheres. Terminations of input and outputs are layer specific. Also depicted is the excitatory interlaminar circuit, whereby information proceeds from granular (white) to supragranular (gray) and then to infragranular (dark gray) layers. Modified from ref. 3.

Table S1. Factor analysis of receptive field parameters using the varimax approach (19)

Parameter	Factor 1	Factor 2	Factor 3	Factor 4
Latency	−0.050	−0.003	0.334	**0.828**
Firing rate	**−0.793**	0.266	0.057	−0.150
Phase locking	**0.859**	0.331	0.076	−0.159
STA separability	**0.647**	0.113	0.218	−0.502
STA feature selectivity	**0.919**	−0.153	0.007	−0.007
First MID nonlinearity asymmetry	0.168	0.198	**0.758**	0.016
MID 2D nonlinearity inseparability	0.411	**−0.471**	−0.452	0.395
First MID contribution	0.133	**0.855**	0.186	−0.206
MID synergy	−0.009	**−0.936**	−0.038	−0.023
Second MID nonlinearity asymmetry	−0.013	0.123	**−0.575**	−0.210
Second MID separability	0.322	**0.654**	−0.382	0.114

The eigenvalues for the analysis were 3.35, 2.35, 1.42, and 1.07, accounting for 74.3% of the data variance. The greatest factor contribution of each parameter is shown in bold.