

Contribution of Inhibition to Stimulus Selectivity in Primary Auditory Cortex of Awake Primates

Srivatsun Sadagopan and Xiaoqin Wang

Laboratory of Auditory Neurophysiology, Departments of Neuroscience and Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Recent studies have demonstrated the high selectivity of neurons in primary auditory cortex (A1) and a highly sparse representation of sounds by the population of A1 neurons in awake animals. However, the underlying receptive field structures that confer high selectivity on A1 neurons are poorly understood. The sharp tuning of A1 neurons' excitatory receptive fields (RFs) provides a partial explanation of the above properties. However, it remains unclear how inhibitory components of RFs contribute to the selectivity of A1 neurons observed in awake animals. To examine the role of the inhibition in sharpening stimulus selectivity, we have quantitatively analyzed stimulus-induced suppressive effects over populations of single neurons in frequency, amplitude, and time in A1 of awake marmosets. In addition to the well documented short-latency side-band suppression elicited by masking tones around the best frequency (BF) of a neuron, we uncovered long-latency suppressions caused by single-tone stimulation. Such long-latency suppressions also included monotonically increasing suppression with sound level both on-BF and off-BF, and persistent suppression lasting up to 100 ms after stimulus offset in a substantial proportion of A1 neurons. The extent of the suppression depended on the shape of a neuron's frequency-response area ("O" or "V" shaped). These findings suggest that the excitatory RF of A1 neurons is cocooned by wide-ranging inhibition that contributes to the high selectivity in A1 neurons' responses to complex stimuli. Population sparseness of the tone-responsive A1 neuron population may also be a consequence of this pervasive inhibition.

Introduction

Neurons in the primary auditory cortex (A1) of awake animals exhibit a high degree of response selectivity, more so than what has been observed in anesthetized animals (Hubel et al., 1959; Evans and Whitfield, 1964; Wang, 2007; Hromádka et al., 2008; Sadagopan and Wang, 2008, 2009). Previous studies have described "narrow" or "well tuned" excitatory responses in awake animals (Suga, 1965; Abeles and Goldstein, 1972; Pelleg-Toiba and Wollberg, 1989; Sadagopan and Wang, 2008) and humans (Bitterman et al., 2008), which partially explains this high selectivity, and consequently the population sparseness of the neural representation of sounds in A1. However, the excitatory areas of the receptive field (eRFs) alone cannot entirely explain such high selectivity, as many stimuli with spectral content overlapping a given neuronal eRF could then drive that neuron. The absence of responses to many such stimuli may be attributed to inhibitory regions of the receptive field (iRFs). We studied iRFs of neurons in A1s of awake marmosets to understand their role in shaping these highly selective responses.

Recordings from A1 of anesthetized animals (Suga, 1965; Arthur et al., 1971; Calford and Semple, 1995; Sutter et al., 1999; Loftus and Sutter, 2001; Sutter and Loftus, 2003) and awake animals (Shamma and Symmes, 1985; Blake and Merzenich, 2002; Barbour and Wang, 2003; O'Connor et al., 2005) have revealed suppression along the frequency axis using two-tone protocols and other types of spectrally dense stimuli. Pharmacological studies suggest that active cortical inhibition might underlie this side-band suppression in bats, causing expansion of RFs, reduction in FM direction selectivity, and degradation of motion processing (Chen and Jen, 2000; Firzlafl and Schuller, 2001; Jen et al., 2002; Razak and Fuzessery, 2009). However, data that characterize inhibitory influences along the sound level and time axes in awake animals are currently lacking.

Earlier studies in anesthetized cats proposed that sound-level tuning could be a two-stage inhibitory process (Calford and Semple, 1995; Sutter and Loftus, 2003). Disproportionate inhibition at high sound levels has been observed intracellularly in a smaller number of neurons in anesthetized cats (Ojima and Murakami, 2002) and anesthetized rats trained using noise stimuli (Tan et al., 2007). Whole-cell recordings from level-tuned neurons suggested that excitation was already tuned to sound level and inhibition increased monotonically with increasing sound level, but within a "nonmonotonic zone" in rat auditory cortex (Wu et al., 2006). In awake animals, however, spike rate-based sound-level tuning is qualitatively different, both in terms of prevalence (Suga, 1965; Brugge and Merzenich, 1973; Pfingst and O'Connor 1981; Sadagopan and Wang, 2008) and distribution throughout A1 (Pfingst and O'Connor, 1981; Sadagopan

Received Oct. 12, 2009; revised March 17, 2010; accepted April 12, 2010.

This work was supported by National Institutes of Health Grant DC-03180 (to X.W.). We thank Dr. Yi Zhou and Dr. Edward Bartlett for helpful comments and suggestions. We also thank Ashley Pistorio and Jenny Estes for assistance with animal care.

Correspondence should be addressed to Xiaoqin Wang, Department of Biomedical Engineering, 720 Rutland Avenue, Traylor 410, Baltimore, MD 21205. E-mail: xiaoqin.wang@jhu.edu.

S. Sadagopan's present address: Department of Neurobiology and Physiology, Northwestern University, Evanston, IL 60208.

DOI:10.1523/JNEUROSCI.5072-09.2010

Copyright © 2010 the authors 0270-6474/10/307314-12\$15.00/0

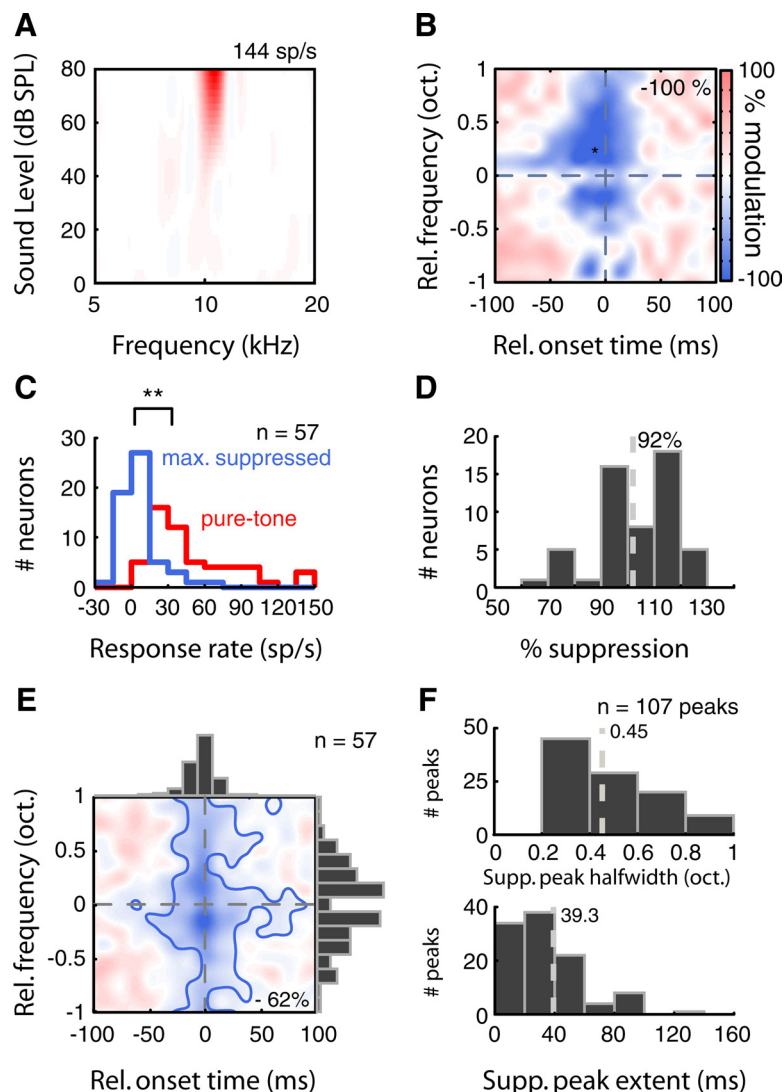


Figure 1. Example of side-band suppression (Supp.) in A1 neurons. **A**, Example FRA of a V unit. This unit was sharply tuned (bandwidth, 0.28 oct. at BL), with monotonically increasing response to sound level. Color map corresponds to the strength of the excitatory response; the maximal response rate of this neuron was 144 spikes/s. **B**, When we probed this unit with two-pip stimuli, we observed strong suppression at higher and lower frequencies that reduced responses to BF/BL pips by up to 100% at maximal suppression (asterisk). **C**, Distributions of pure-tone (red histogram) and maximally suppressed two-pip (blue histogram) firing rates of 57 neurons. $^{**}p < 0.01$, Wilcoxon rank-sum test. **D**, Distribution of percentage suppression observed over the population of 57 neurons. On average, we observed 92% maximal suppression by an off-BF pip (gray dashed line is mean). **E**, The population average of two-pip suppression maps from 57 neurons. Blue shading corresponds to the percentage of suppression, and blue contour indicates significance at $p < 0.05$ level (t test). Histograms on the margins are locations of peak suppression in frequency and time; we did not observe any systematic asymmetries in the location of suppressive peaks. **F**, Distribution of the bandwidth of suppressive peaks (top; $n = 107$ peaks; gray dashed line is median, 0.45 octaves). At half-maximal suppression, each peak had a bandwidth ~ 1.7 times the bandwidth of the excitatory peak measured using single tones. At half-maximal level, the median temporal extent of suppressive peaks was 39.3 ms (bottom). In 50 of 57 neurons, we observed two suppressive regions located above and below BF. Rel., Relative.

and Wang, 2008). Therefore, it is necessary to characterize the suppressive effects of sound level in an awake preparation.

Here, in addition to two-tone suppression, we measured single-tone suppression (stSUP), defined as the suppression of firing rate below spontaneous rate in response to a single-tone stimulus, along three axes important to audition, namely, frequency, sound level, and time in A1 neurons of awake marmosets. We found strong suppression at side-band frequencies and strong stSUP that increased monotonically with sound level both on- and off-BF in a large population of A1 neurons, with effects lasting for at least 100 ms after stimulus offset. Based on these

data, we conclude that finely tuned excitation in frequency and sound level is cocooned by broadly tuned inhibition in frequency, sound level, and time in a large proportion of neurons in A1 of awake marmosets. Our results suggest that A1 neurons in awake animals are under the influence of much more pervasive inhibition than previously described, which may contribute to their highly restrictive sampling of acoustic stimulus space (Wang, 2007).

Materials and Methods

Neural recording

We recorded single-unit (SU) responses from A1s of three awake marmoset monkeys. All experimental procedures were in compliance with the guidelines of the National Institutes of Health and approved by the Johns Hopkins University Animal Care and Use Committee. A typical recording session lasted 4–5 h, during which an animal sat quietly in a specially adapted primate chair with its head immobilized. The experimenter ensured that the animal's eyes were open before presentation of a given stimulus set. Technical details of surgical and experimental procedures are described in a previous publication (Lu et al., 2001). Briefly, a tungsten microelectrode (A-M Systems; impedance, 2–4 M Ω) was positioned within a small craniotomy (~ 1 mm diameter) using a micromanipulator (Narishige Instruments) and advanced through the dura into cortex using a hydraulic microdrive (Trent Wells). The experimenter typically advanced the electrode by ~ 25 μ m and waited for a few minutes to allow the tissue to settle. During this period, a set of search stimuli were played that typically consisted of pure tones [approximately five steps per octave (oct.)], bandpassed noise, and marmoset vocalizations at multiple sound levels. This strategy of “burst” electrode movements with long waits while playing a wide array of stimuli helped us detect and isolate single units with very low spontaneous activity and avoid biases toward any particular kind of units. Proximity to the lateral sulcus, clear tone-driven responses in the middle cortical layers, and a tonotopic relationship with other recorded units were used to determine whether a recording location was within A1. All reported data are based on well isolated single units. Single-unit discrimination was performed online using a template match algorithm (Alpha-Omega). All data were collected from marmosets listening to sounds in an awake, passive state. We sampled over a large range of best frequencies (BFs) (0.5–20 kHz) and best levels (BLs) [0–80 dB sound pressure level (SPL)].

Acoustic stimuli

Stimuli were generated digitally in MATLAB (MathWorks) at a sampling rate of 97.7 kHz, converted to analog signals (Tucker-Davis Technologies), attenuated (Tucker-Davis Technologies), power amplified (Crown Audio), and played from a loudspeaker (FT-28D or B&W-600S3, Fostex) situated ~ 1 m in front of the animal. The loudspeaker had a flat frequency-response curve (± 5 dB) across the range of frequencies of the

stimuli used, with a calibrated level (at 1 kHz) of ~ 90 dB SPL at a set level of 0 dB attenuation. Stimuli consisted of pure tones, usually 100 ms long with 5 ms cosine ramps, delivered every 300 ms in pseudo-random order, and repeated five to eight times. We sampled the frequency and level axes with sufficient density so that at least two significant response bins could be expected in most cases. For two-pip suppression experiments, we used two short (20–40 ms) tone pips, one centered at the BF of the neuron being tested and the other roved in both frequency and onset time with respect to the BF pip. We sampled the frequency axes every one-eighth octave, and the time axis in 12.5 ms bins. The sound levels of both pips were held constant at the BL of the neuron being recorded from. We required at least eight repetitions to be completed to include a neuron in this analysis to ensure statistical significance. We typically centered IFM stimuli at BF and BL, varying both sweep bandwidth and sweep length, resulting in extensive sampling of sweep velocities from ~ 1 to ~ 200 octaves per second.

Data Analysis

Single-tone frequency-response areas. The single-tone frequency-response areas (FRAs) described in this article were derived from the 275 units recorded from the A1 of two awake marmosets. Units were quantitatively classified as “O” ($n = 175$, 64%) or “V” ($n = 100$, 36%) based on FRA shape. The centroid of the frequency-tuning curve at BL was used as an estimate of BF. These eRF properties were described and characterized in an earlier study (Sadagopan and Wang, 2008). To compute population averages, we centered all FRAs at BF and BL in the case of O units, and at BF and threshold (defined as the level that causes at least a 20% response in 20 dB bins) for V units. In most cases, we systematically sampled a ± 1 octave range around BF in 0.1 oct. steps, but this was not always fixed for all the neurons tested. Therefore, in our population averages some frequency bins contained the contributions of a smaller number of neurons, especially near the edges of the sampled frequency range. Similarly, because we only sampled up to a maximum sound level of 80 dB (20 dB steps), some bins in the population average at the loudest sound levels also contained contributions from fewer units. This nonuniformity in sampling is accounted for in all presented averages and statistics. For completeness, the maximum and minimum sampled bins in the population averages displayed in this article are: O maximum: $n = 175$; relative frequency (rel. freq.), 0 octaves; rel. level, 0 dB; O minimum: $n = 42$; rel. freq., -1 octaves; rel. level, 80 dB; V maximum: $n = 100$; rel. freq., 0 octaves; rel. level, 80 dB; V minimum: $n = 14$; rel. freq., -1 octaves; rel. level, -60 dB. All FRAs were normalized to their individual maximal tone response rates before averaging. Colormaps of population FRAs are plotted on a normalized color scale, where deepest red corresponds to $+1$ (normalized maximal firing rate), white corresponds to zero (baseline), and deepest blue corresponds to -1 .

Two-pip suppression maps. Consistent with earlier terminology, we use the word “suppression” to refer to the reduction of responses to a BF/BL tone due to the presence of a second tone. Two-pip suppression maps were computed in every time-frequency bin as, $\% \text{Suppression} = (R_{2\text{pip}} - R_{\text{BF}}) / R_{\text{BF}} * 100$, where $R_{2\text{pip}}$ is the combination response rate after subtraction of spontaneous rate and R_{BF} is the response at best frequency after subtraction of spontaneous rate. Note that suppression exceeding 100% can result if the combination response results in suppression of spiking below the spontaneous rate. Two-pip data were collected from 91 units on which at least eight repetitions of every frequency/time combination were completed. We used a t test to compare the two-pip response in the maximally suppressed bin to the average BF pip response, requiring significant suppression at $p < 0.05$ (t test) for inclusion in further analyses. Fifty-seven of ninety-one (62.6%) units met this criterion, showing a significant suppressive effect attributable to the non-BF pip. The population average of the two-pip map was computed using all 57 neurons in every bin; as in this case, all bins were uniformly sampled across the population. All averages were performed after normalization.

IFM responses. We sufficiently sampled the range of IFM sweep velocities from ~ 1 to ~ 200 oct./s in 51 of the 57 neurons that showed significant two-pip suppression. In many cases, we used IFM sweeps roving in both bandwidth and length, resulting in dense sampling of intermediate

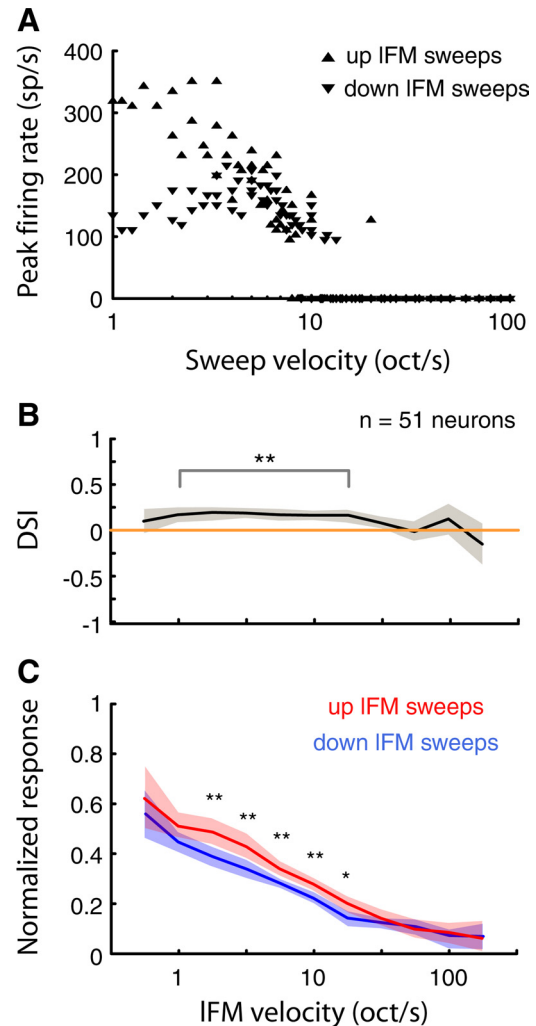


Figure 2. Consequence of side-band inhibition for the tone-contour processing. **A**, Tuning of the neuron in Figure 1A in response to upward and downward IFM sweeps of different velocities clearly exhibiting a low-pass characteristic. **B**, **C**, Over the population, we found that both upward and downward IFM sweeps were strongly low passed, consistent with the average suppression map in Figure 1E. Red line and shading are the mean and the 95% confidence interval for the upward sweep population; blue line and shading are corresponding quantities for downward sweeps. Unexpectedly, we observed higher rates for upward sweeps over a range of intermediate velocities (**B**, black line is mean population DSI; shading is the 95% confidence interval). $**p < 0.01$, $*p < 0.05$, two-tailed t test for **B** (comparing DSI at each velocity to zero), and paired t test for **C** (pairwise comparison of up- and down-IFM responses at each velocity). oct, Octave.

velocities. We did not find any systematic effects of sensitivity along the individual bandwidth or length dimensions that could not be explained in terms of velocity for the tested parameter ranges; therefore, for population averaging we restricted our analyses to velocity tuning curves. Peak firing rates in response to each stimulus were calculated using non-overlapping 20 ms bins. Each stimulus set consisted of both up- and down-IFM sweeps, and normalization was based on the sweep that elicited the highest firing rate regardless of direction. A direction-selectivity index (DSI), defined as the difference between the peak upward and downward sweep responses divided by the sum, was used to test whether neurons preferred upward (DSI > 0) or downward (DSI < 0) sweeps.

Response model. From our data, we observed that the spike count distribution of a given neuron responding to a stimulus set in 50 ms time bins was asymmetric with a heavy right tail, consistent with a log-normal distribution. Therefore, we modeled each neuron’s response distribution to a stimulus set arising from a log-normal distribution. The mean of the distribution was held constant at 30 spikes/s, or 1.5 spikes

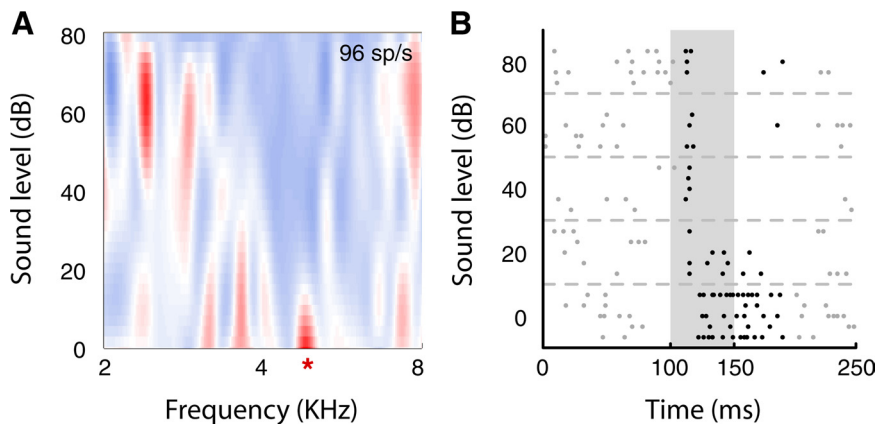


Figure 3. Single-unit example showing widely tuned inhibition on- and off-BF at loud sound levels. **A**, The FRA of an example neuron with high spontaneous activity constructed by presenting five repetitions of 155 pure tones of varying frequency and sound level. This neuron's excitatory response peaked at 4.8 kHz per 0 dB SPL (red asterisk), with a maximal response rate of 96 spikes/s. Widely tuned inhibitory regions (blue regions) were evident both on- and off-BF at louder sound levels. **B**, Spike raster of this neuron to five tested sound levels at BF. Spike responses are clearly diminished at sound levels louder than the BL of this unit (0 dB) at BF. Shaded region corresponds to stimulus duration. Gray and black dots correspond to spontaneous spikes and spikes falling within the analysis window, respectively.

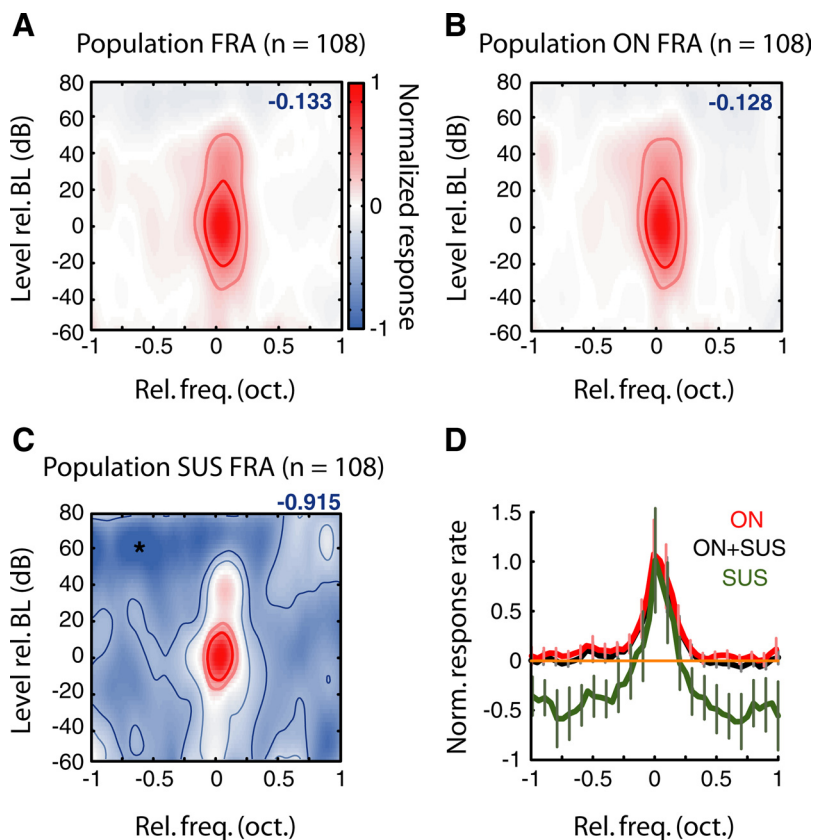


Figure 4. Side-band stSUP can be revealed by population averaging. Population averages of the sharpest 108 (of 275) units centered on BF and BL are plotted. Color map corresponds to normalized response rate; dark and light red contours correspond to 50 and 25% of maximal excitation; dark and light blue lines are corresponding quantities for stSUP. **A**, **B**, No stSUP was apparent when the responses were computed over the entire response duration or only during the onset portion of the response. **C**, When averaged only during the sustained portion of the response (window from 50 ms after stimulus onset to 50 ms after stimulus offset, 100 ms stimuli), strong stSUP is evident (number is inhibitory strength relative to normalized excitatory strength). Asterisk denotes maximally inhibited bin. **D**, Average frequency-tuning curves for these 108 units computed at BL, for the entire response duration (black), onset response (red), and sustained response (green). Orange line corresponds to baseline; error bars denote the 95% confidence interval.

in each modeled 50 ms bin. The kurtosis (K) of a log-normal distribution is a monotonically increasing function of its shape parameter (σ) as follows:

$$K = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3.$$

Using this relationship, for the experimentally obtained kurtosis values at given time points, we obtained the corresponding shape parameter graphically. We simulated 300 neurons responding to a set of 300 stimuli by drawing paired kurtosis values at the relevant time points, with resampling, from our population of 175 O units.

Simulation. We simulated the effect of recording multiunit (MU) activity of many single units by summing n FRAs, with replacement, from the set of 275 FRAs we had collected (64% O/36% V mix), where n was the number of single units contributing to a given simulated multiunit signal. This allowed us to systematically study the effect of multiunit quality. In this case, it is important to note that we did not normalize firing rates before summing the FRAs. Because A1 is topographically organized, we assumed that a given electrode penetration sampling from nearby units is likely to record from units with similar BFs. Therefore, all FRAs were centered at BF, with frequency tuning being relative to BF. However, we found no evidence for a systematic cortical organization of neurons by BL (see Fig. 9), and, therefore, the sound-level axis was absolute. For each n , we simulated recording from 200 multiunit "sites," each containing a randomly drawn summation of n single units. We computed the monotonicity index (MI) as the ratio of the response at the loudest level to the response at the BL; MI was 0 for highly non-monotonic units and 1 for monotonic units. A given simulated unit was classified as non-monotonic for $MI \leq 0.5$.

Results

Analyses presented here are based on single-tone responses from 275 well isolated single units and two-pip responses from 57 single units, some overlapping with the earlier population. Based on an earlier study, we classified single-tone responses into O ($n = 175$, 64%) or "I/V" ($n = 100$, 36%) units depending on the shape of their FRAs (Sadagopan and Wang, 2008). We recorded two-pip responses from both O ($n = 51$) and narrowly tuned I/V ($n = 6$) units but did not separate them in further analyses as there were no systematic differences in their side-band suppression maps. The inhibitory effects are analyzed and presented below along three axes, frequency, sound level, and time.

Side-band suppression around BF at BL
First, we analyzed the effects of the presence of a second tone pip on the responses

to a tone pip at BF and BL. Figure 1A plots the FRA of an example of an I/V unit, which was sharply tuned in frequency (tuning bandwidth at 80 dB, 0.28 oct.) and had a monotonic response to sound level. When we plotted the suppression map of this neuron in frequency and time (Fig. 1B), we observed strong suppression of the response to the BF tone (100% suppressed) at both higher and lower frequencies. Over a population of 57 tone-responsive neurons (median tone response at BF/BL, 32.5 spikes/s; spontaneous rate subtracted) (Fig. 1C, red histogram), the addition of a second off-BF pip resulted in significantly suppressed responses (median maximally suppressed response, 1.4 spikes/s; spontaneous rate subtracted) (Fig. 1C, blue histogram). Over the population, the addition of a specific non-BF tone pip suppressed the BF tone's response by 92% on average (Fig. 1D). Importantly, in 16 of 57 neurons (28%), we observed >90% suppression one-eighth oct. both above and below BF. The effective bandwidths of these neurons were therefore on the order of half of our sampling resolution of the frequency axis ($\sim 1/16$ oct.). It should be noted that this suppression occurred at BL, where excitatory frequency tuning is typically at its widest for O units.

To better understand the nature of side-band suppression over the population of tone-responsive neurons, we averaged all 57 two-pip suppression maps, centered on the BF pip. The resultant average has been interpolated and smoothed for visualization (Fig. 1E). In most neurons (50 of 57 neurons), we observed two significant suppression peaks at both higher and lower frequencies. The locations of the suppressive peaks were symmetrically distributed about BF on average (Fig. 1E, marginal histogram on y -axis). Significant suppression lasted for up to ~ 30 ms after stimulus offset. In the ($\Delta f = 0$, $\Delta t = 0$) trial condition, where f is frequency and t is time, an average suppression of 31% was observed. In this trial condition, the actual stimulus presented consisted of two in-phase, synchronous tone pips at the same frequency that were summed, resulting in an amplitude 6 dB higher than that in the single-pip condition. For some strongly nonmonotonic units, this increase in sound level was sufficient to result in a modest suppression of response. The median bandwidth (at 50% suppression) of each peak was 0.45 oct. (Fig. 1F, top), about 1.7 times of the observed median excitatory peak bandwidth (0.27 oct.). Similarly, median peak extent of each suppressive peak was ~ 40 ms (Fig. 1F, bottom), comparable to a study in awake owl monkeys using spectrotemporal receptive field methods (Blake and Merzenich, 2002). The duration of suppression was longer at frequencies that were closer to BF. The population average half-maximal temporal extent of the inhibitory region at ± 0.125 oct. from BF was 34 ms, whereas at ± 0.5 oct., inhibitory effects lasted 12.5 ms. However, both the maximal degree

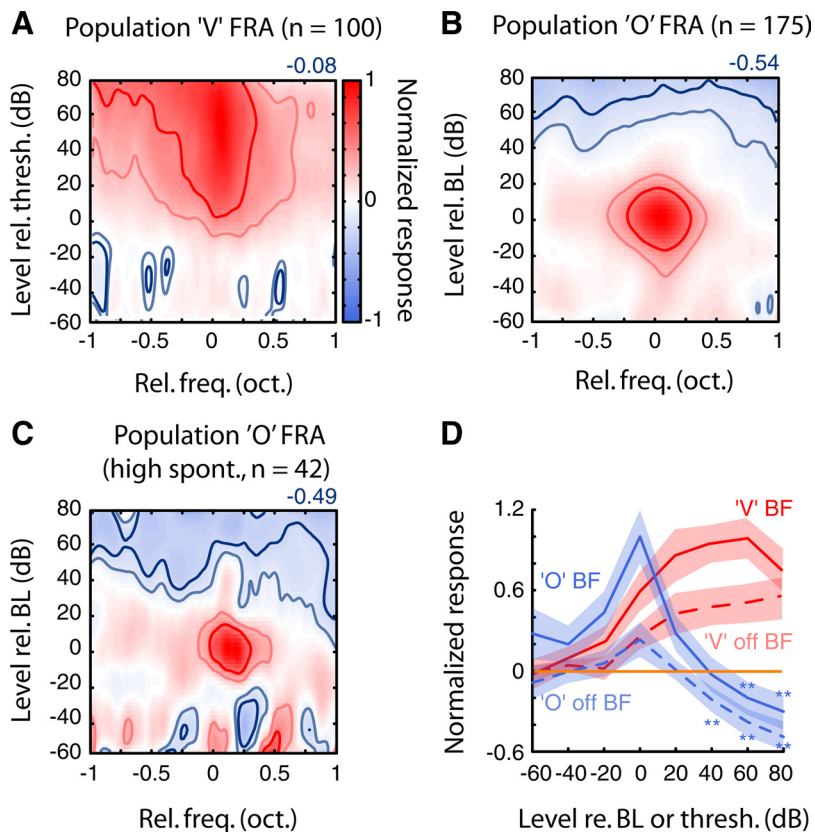


Figure 5. Population O and V unit FRAs show differential effects of stSUP. **A**, The population average FRA of $n = 100$ V units centered on BF and threshold. Color map and contours are as described earlier. For V units, stSUP was only -8% of maximal excitation in this range. The image has been interpolated and smoothed for display purposes. **B**, Similar population average of $n = 175$ O units, but centered at BF and BL. Monotonically increasing inhibition, over a wide frequency range enveloping BF is apparent at loud levels. Maximal inhibitory strength was -54% of the maximum excitation. **C**, Same as **B**, but for $n = 42$ O units that exhibited a high spontaneous rate (>4 spikes/s). Inhibition at loud levels is even more apparent. **D**, Population average rate-level curves on-BF (solid lines) and off-BF (dashed lines) for O (blue) and V (red) units showing differential action of stSUP. Lines correspond to population means, shadings correspond to the 95% confidence interval, line and shading colors are as indicated. For O units, the inhibition observed both on- and off-BF at loud levels was statistically significant ($**p < 0.01$, two-tailed t test comparing normalized rate distribution at each sound level above BL showing suppression with baseline). thresh., Threshold.

of suppression and the duration of suppression did not depend on the length of the tone pips used (20–40 ms).

We hypothesized that the temporal extent of the suppression at both higher and lower frequencies near BF, evident from both the single-neuron and average suppression maps, would impart a low-pass characteristic on the processing of tonal contours. For example, a fast ILM sweep would pass through suppressive regions of the receptive field in both directions (up and down) before crossing the excitatory region of the receptive field, resulting in severe suppression of the neuron's spiking response to the sweep. However, if the sweep were slow enough, it would avoid the inhibitory regions, resulting in a strong spiking response. When we tested these scenarios in the example unit of Figure 1A using ILM sweeps of varying bandwidths and lengths (therefore of different velocities), we observed a clear low-pass characteristic in the velocity-response profile in both directions, consistent with our predictions (Fig. 2A). Note that we used a particularly dense sampling only for this neuron, which is not representative of the remainder of the population. When we plotted the population response-velocity response profiles of 51 neurons tested using ILM sweeps of varying velocities, we observed that ILM responses were indeed clearly low pass for both up and down directions (Fig. 2B,C). Unexpectedly, we observed a small but

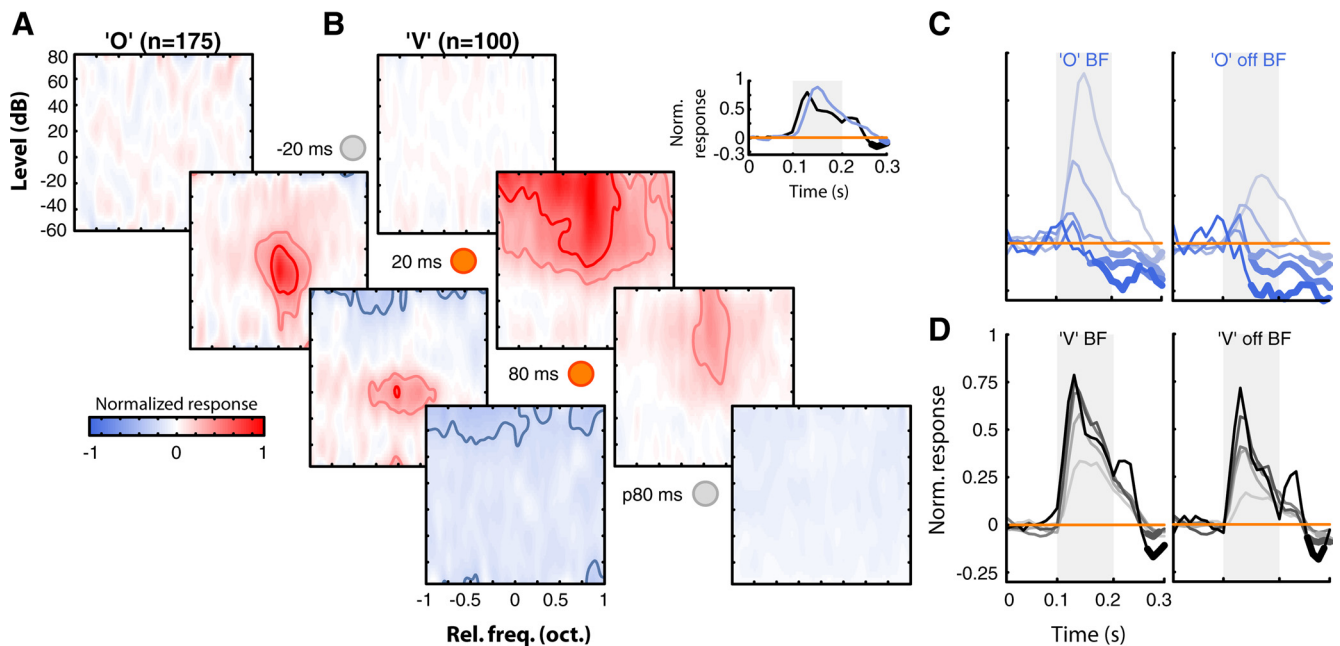


Figure 6. Temporal dynamics of O and V unit population responses. **A**, Population FRAs of O units in 20 ms bins starting at four time points: (1) 20 ms before stimulus onset; (2) 20 ms after stimulus onset; (3) 80 ms after stimulus onset; and (4) 80 ms poststimulus offset. Color map and contours are as described earlier. Gray and orange discs indicate stimulus off and on, respectively. Early excitatory level tuning and late development of stSUP at loud sound levels is apparent. **B**, Similar profile for the V unit population average FRA. No long-lasting stSUP could be observed. V unit excitatory responses developed earlier than O unit responses and adapted more during response duration (inset; black line is V unit population PSTH at BF/BL, blue line is O unit population PSTH at BF/BL, gray shading is stimulus duration). **C**, Population PSTHs of O units on- and off-BF. Intensity of line color corresponds to increase in sound level in 20 dB steps from BL, lightest blue line is PSTH at BL. Thicker parts of the line below zero are significantly suppressed ($p < 0.05$, t test). stSUP develops late and persists up to at least 100 ms after stimulus offset. **D**, Corresponding PSTHs for V units. Intensity of line color corresponds to increasing sound level, lightest gray line is PSTH at 0 dB SPL, black line is PSTH at 80 dB SPL. Thicker parts of the line below zero are significantly suppressed ($p < 0.05$, t test). No long-lasting stSUP is apparent in this case.

significant response bias for upward IFM sweeps (Fig. 2*B*) as measured using a DSI (see Materials and Methods). We could not explain this selectivity for upward sweeps in terms of the two-pip maps obtained earlier, as we did not observe any systematic asymmetry in the extents of suppressive peaks. The range of velocities over which this differential preference was observed was similar to an earlier study in anesthetized rats (Zhang et al., 2003), but our observed DSIs were much lower. However, important differences between our data and those of Zhang et al. (2003) should be noted here: (1) neither the DSI nor the magnitude of inhibition systematically varied with BF in our data; and (2) most of the units we recorded from were O units with fine frequency and sound-level tuning as opposed to mostly V units in the Zhang et al. (2003) study. Therefore, it is unclear whether an “edge effect,” similar to the one described by Zhang et al. (2003), caused this differential upward sweep selectivity. Some marmoset vocalizations contain FM contours that overlie this range of IFM velocities (Agamaite, 1997; DiMattina and Wang, 2006), but it is as yet unclear whether the observed phenomenon could underlie preferential processing of particular vocalizations reported earlier (Wang and Kadia, 2001).

stSUP around BF over a wide range of sound levels

To better relate our data to available intracellular recordings from anesthetized animals, we needed a way to characterize inhibition using single tones. In general, using single-tone paradigms, significant stSUP (firing rate decreased below spontaneous rate) was rarely observed at the single-unit level using conventional extracellular recording techniques. This was attributable to the low spontaneous rates of neurons we typically observed in our recorded population (median: O units, 1.74 spikes/s; V units, 4

spikes/s). However, in a few units with high spontaneous rates, we could observe significant suppression of firing rates below spontaneous rates that were consequences of inhibitory inputs to these neurons. For example, the nonmonotonic neuron in Figure 3 showed clear suppressive regions around BF (blue areas in the FRA). A closer inspection of the spike raster over different sound levels at BF revealed a lack of spiking at loud sound levels (Fig. 3*B*).

Over the population of recorded neurons, this underlying stSUP could also be revealed by computing population average FRAs of a large number of neurons ($n = 175$ O units and $n = 100$ V units). In Figure 4, we present population averages of the sharpest 108 of 275 units that were responsive to pure tones (both O and V shaped), consisting of those units that exhibited bandwidths less than the mean bandwidth of the entire population. A population-averaged FRA was obtained by aligning individual FRAs at the BF and BL of each unit. When the population average FRA was computed from the entire response duration (Fig. 4*A*) or only the onset portion of the response (Fig. 4*B*, “ON”), we did not observe any suppression of firing rate below the spontaneous rate. However, the population average FRA computed during the sustained portion of the response (50 ms after stimulus onset to 50 ms after stimulus offset) (Fig. 4*B*, “SUS”) revealed strong stSUP (Fig. 4*C*). The blue areas in Figure 4*C* indicate persistent stSUP at all sound levels at side-band frequencies and at levels higher than BL at BF. Note that strong stSUP is present even at low intensities of off-BF tones, an observation consistent with an earlier intracellular study of A1 neurons in anesthetized rats using two-tone stimuli (Scholl et al., 2008). Figure 4*D* shows a “slice” of the average FRA at BL. During the sustained response, the bandwidth of the excitatory response was narrower compared with

that of the onset response. Note that we chose the sharpest 108 units for this analysis to uncover side-band stSUP specifically; had we used all 275 units, strong excitatory responses of widely tuned neurons would have averaged out the observed inhibitory effect.

On- and off-BF stSUP at loud sound levels

Figure 5A shows the population average FRA of the V units, aligned at BF and threshold sound level (defined as the sound level evoking 20% response). The FRAs were, on average, asymmetric with a long low-frequency tail. No stSUP was observable at BF or within ± 1 oct. of BF. However, when we plotted the population average of O units aligned at BF and BL, we observed strong stSUP at loud sound levels both on- and off-BF, over the entire two-octave range of sampled frequencies (Fig. 5B). This stSUP was especially apparent when we only averaged across O units with a high spontaneous rate (Fig. 5C) ($n = 42$ of 175 O units with a spontaneous rate of >4 spikes/s), as this enabled us to better observe a larger decrease in firing rate from an extracellular perspective. All these population FRAs were computed over the entire stimulus duration. The maximal observed inhibitory magnitude was -54% of normalized excitatory response in O units, and the observed stSUP was statistically significant ($p < 0.01$, t test) at sound levels that were at least 40 dB (two sampling bins) louder than BL. Population average, rate-level curves for O (blue) and V (red) units at BF (solid lines) and off-BF (0.5 octaves below BF for illustration purposes; dashed lines) are plotted in Figure 5D. V unit responses were generally monotonically increasing with sound level both on- and off-BF. O units, however, showed significant ($p < 0.01$, t test) stSUP at sound levels that were louder than BL, with off-BF stSUP being stronger than on-BF stSUP. These results suggest that inhibition might overwhelm excitation at loud sound levels, as proposed by Ojima and Murakami (2002) in an earlier intracellular study in anesthetized cats. However, we faced the same interpretative limitations as in that study; because we could only observe the net effect of excitation and inhibition, we could not conclude whether the inhibition we observed was causally responsible for sound-level tuning.

On- and off-BF stSUP over time

To address the limitation mentioned above, we constructed population average FRAs in 20 ms time bins to see whether we could observe progressive shape changes to O unit FRAs over the response duration. Figure 6A and B, shows “snapshots” of the population FRA at four time points: (1) before stimulus onset; (2) 20 ms after stimulus onset (to account for latency); (3) 80 ms after stimulus onset; and (4) 80 ms after stimulus offset. Consistent with our earlier observation (Sadagopan and Wang, 2008), the population V FRA was initially broadly tuned, becoming more

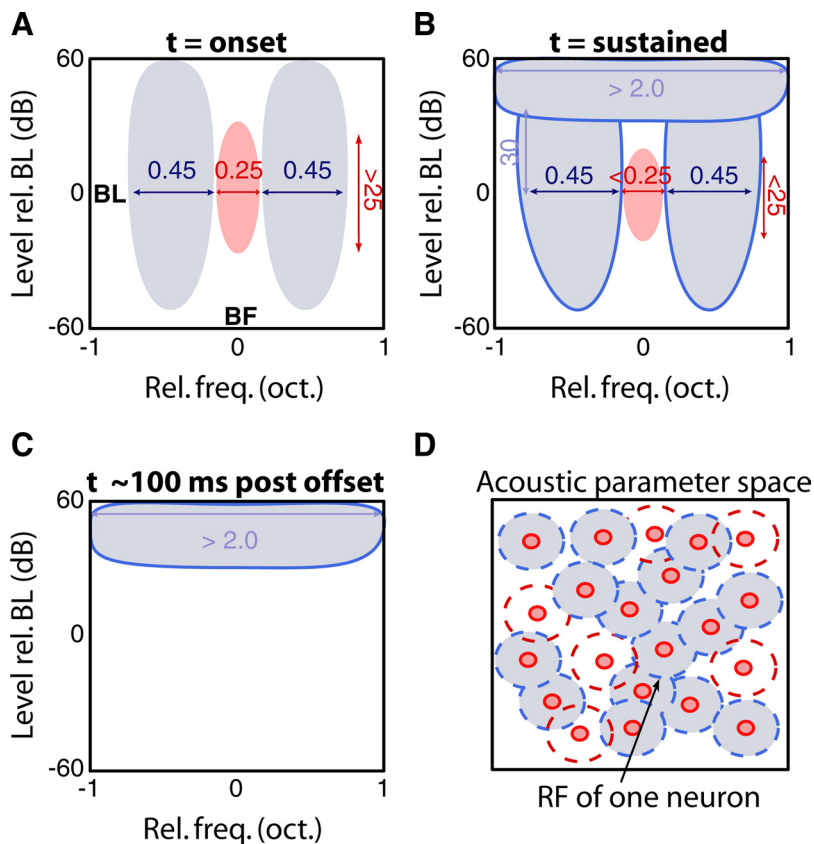


Figure 7. Schematic of stSUP effects on O units. **A**, At stimulus onset, excitation (red) is tuned to level and frequency, flanked by suppressive side bands (blue) that are ~ 3.5 times as wide as the excitatory peak. **B**, During the sustained phase of the response, a late stSUP component develops at loud sound levels, widely tuned in frequency. This late stSUP, together with the suppressive side bands, effectively envelopes the excitatory part of the receptive field. During this phase, frequency and level tuning of the excitatory component may be sharpened. **C**, While the side-band suppression disappears soon after stimulus offset, the stSUP at loud levels remains in effect for at least 100 ms after stimulus offset (tested range). **D**, These inhibitory effects effectively create regions of inhibition that surround the excitatory regions during the sustained response in a large population of A1 units (O units). The V units do not exhibit inhibitory effects but show narrowing of the excitatory receptive field as response progresses. Adapted from Wang (2007).

narrowly tuned and O-like as the response progressed. In contrast, the O units exhibited strong sound-level tuning throughout the response duration. Importantly, the O units were already tuned to the level (albeit more broadly) in the first significant response bin (20 ms after stimulus onset), suggesting that the excitatory inputs to these units are themselves perhaps level tuned. These data still did not conclusively inform us whether level tuning is inherited. For example, it has been suggested in anesthetized animals that an early inhibitory component might shape excitatory tuning to sound level (Calford and Semple, 1995; Sutter and Loftus, 2003). However, in contrast to those studies, we also observed a late inhibitory component that was strongly monotonic and broadly tuned in frequency and further shaped the FRA (Fig. 7A, at 80 ms after stimulus onset and 80 ms after stimulus offset).

When we analyzed the population peri-stimulus time histograms (PSTHs) of O units (Fig. 6C), we observed that the effects of stSUP at loud sound levels, both on- and off-BF, typically started >30 ms after stimulus onset and reached maximal inhibition at >50 ms after stimulus onset. These effects persisted for at least 100 ms after stimulus offset (our recording window for most of these neurons), showing no trend of returning to baseline within this period. This prolonged stSUP was statistically significant (thicker lines below zero meet a criterion of $p < 0.05$, t test)

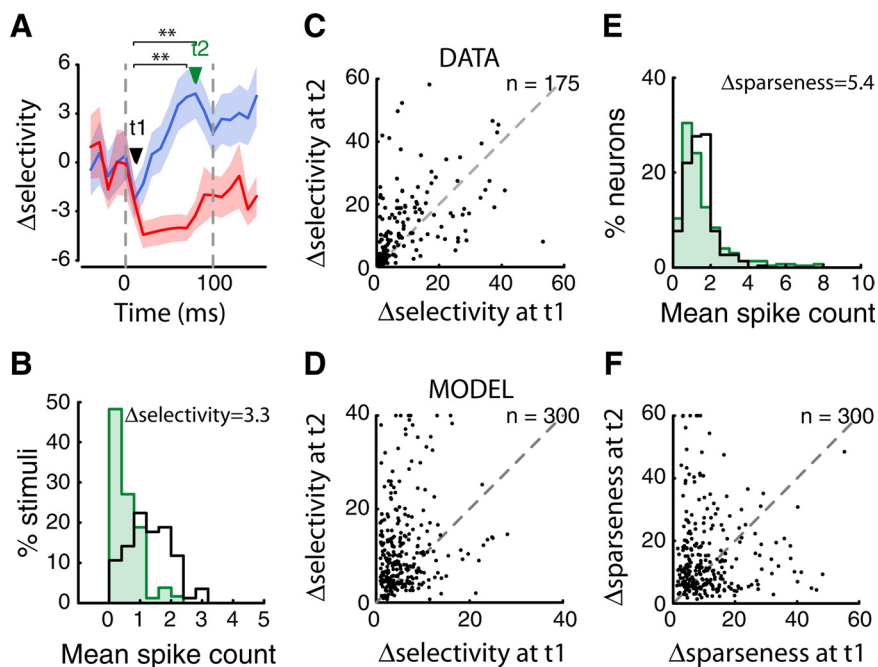


Figure 8. Contribution of stSUP to stimulus selectivity and population sparseness. *A*, Selectivities of O (blue) and V (red) units as a function of time, calculated using 50 ms overlapping bins (10 ms steps). Mean (lines) and 95% confidence intervals (shading) are plotted. Dashed gray lines correspond to stimulus duration. t1 (black arrow) and t2 (green arrow) are time points over which O unit selectivity showed the greatest increase. $**p < 0.01$, multiple-comparison corrected ANOVA. *B*, Spike count distributions of an example neuron responding to 105 tone stimuli at times corresponding to t1 (black) and t2 (green), evaluated from 50 ms bins centered at these times. *C*, Comparison of selectivities evaluated at times t1 and t2 for $n = 175$ O units. *D*, A simple model of A1 responses could qualitatively capture the selectivity increase observed in real data. *E*, Distribution of the responses of a population of 300 model neurons evoked by a single stimulus at times t1 and t2, evaluated in 50 ms bins. *F*, Comparison of the population sparseness of 300 model neurons responding to 300 stimuli at times t1 and t2.

and was observable later in the response at all sound levels including BL. In contrast, we did not observe such prolonged stSUP for V units (Fig. 6*D*); typically, loud sound levels drove these units to higher firing rates. In fact, at the loudest levels tested, excitatory responses to stimulus offset were evident both on- and off-BF. Some significant stSUP could be observed after the offset response of V units, but this stSUP was short-lived and typically returned to baseline within our recording window.

Summary of suppressive and stSUP effects

The above data can be summarized in the two types of excitatory and inhibitory properties exhibited by O and V units. For the eRF, we observed: (1) O unit eRFs that were tuned to frequency and level at the beginning of their response, but did not substantially sharpen in either dimension as the response progressed; and (2) V unit eRFs that were broadly tuned in frequency with monotonically increasing response to level at the beginning of their response, but showed sharpening in both frequency and level tuning, and became more O-like as the response progressed. As described in our earlier study, during the sustained portion of the response $\sim 76\%$ of the units are O shaped (Sadagopan and Wang, 2008). These eRFs were further shaped by iRFs; the three types of stSUPs we observed can be summarized as follows: (1) two-tone suppression coincident with excitation in both O and V units, broadly tuned in frequency (up to ± 1 octaves from BF, total iRF bandwidth at 50% suppression being ~ 3.5 times eRF bandwidth, with iRF suppressing BF responses by 92% on average), lasting up to ~ 40 ms after stimulus offset (Fig. 7*A*); (2) side-band stSUP that develops during the sustained portion of the response, which further sharpens response tuning (Fig. 7*B*); and (3) monotonically

increasing stSUP only in O units. The stSUP was broadly tuned in frequency (> 2 octaves, iRF bandwidth at loudest level was at least eight times the eRF tuning bandwidth at BL, iRF magnitude being up to 50% of BF/BL tone response), starting late (peaking ~ 50 ms after stimulus onset) and persisting well after stimulus offset (> 100 ms) (Fig. 7*B, C*).

Contribution of stSUP to stimulus selectivity

To correlate how the above stSUP effects shape selectivity, we analyzed how the stimulus selectivity of individual neurons to tone stimuli evolved over stimulus duration. Here, we defined selectivity as the reduced kurtosis of the spike count distribution (Lehky et al., 2005) of the neuron responding to ~ 100 tone stimuli at varying frequencies and sound levels. We measured selectivity in 50 ms overlapping bins (10 ms steps). The mean selectivity of our population of O ($n = 175$; blue) and V ($n = 100$; red) units as a function of time are plotted in Figure 8*A*. For O units, we observed a steady increase in selectivity from stimulus onset, peaking shortly before stimulus offset, a time course that paralleled the evolution of stSUP observed in these units (Fig. 6*C*). At its peak value (Fig. 8*A*, time t2), the selectivity increased by 47% compared with the selectivity immediately following stimulus onset (Fig. 8*A*, time t1) and was statistically significant ($p < 0.001$, multiple-comparison, corrected ANOVA). V units, which did not show an increase in stSUP, did not show a corresponding increase in selectivity over time. Spike count distributions for an example O unit at the maximally divergent time points (Fig. 8*A*, t1 and t2) are plotted in Figure 8*B*. Over the population of O units, selectivity at time t2 was consistently greater than selectivity immediately following stimulus onset (Fig. 8*C*). O units that exhibit significant responses during both the onset and sustained portions of the response do not show significant changes in eRF areas over time (Sadagopan and Wang, 2008). Therefore, a shift of the response distribution toward zero, caused by inhibition at nonpreferred stimuli, could be a primary factor in causing increased selectivity.

Population sparseness and its relationship to the observed stSUP, however, could not be estimated from these single-unit data because we did not present identical stimuli to a large population of neurons. However, to infer population sparseness from available data, we constructed a simple statistical model of neural responses in A1. Each neuron's responses to a set of stimuli were modeled to be drawn from a log-normal distribution. We simulated the responses of 300 model neurons responding to 300 stimuli at two time points (t1 and t2). The mean response rate was held constant at 30 spikes/s (1.5 spikes in a 50 ms window), consistent with the physiological response rates of O units. The shape of the distribution at two times corresponding to t1 and t2, which was the only free parameter of the model, was matched to the kurtosis of the data distributions at t1 and t2 (see Materials and Methods). Using this simple model, we were able to qualitatively replicate the pattern of the change in selectivity over time ob-

erved. Population sparseness and its relationship to the observed stSUP, however, could not be estimated from these single-unit data because we did not present identical stimuli to a large population of neurons. However, to infer population sparseness from available data, we constructed a simple statistical model of neural responses in A1. Each neuron's responses to a set of stimuli were modeled to be drawn from a log-normal distribution. We simulated the responses of 300 model neurons responding to 300 stimuli at two time points (t1 and t2). The mean response rate was held constant at 30 spikes/s (1.5 spikes in a 50 ms window), consistent with the physiological response rates of O units. The shape of the distribution at two times corresponding to t1 and t2, which was the only free parameter of the model, was matched to the kurtosis of the data distributions at t1 and t2 (see Materials and Methods). Using this simple model, we were able to qualitatively replicate the pattern of the change in selectivity over time ob-

served in our data (61% increase at t2 compared with t1, $p < 0.001$, paired t test) (Fig. 8D). From this model, we could then obtain the response distribution of the 300 model neurons to each stimulus, from which population sparseness could be estimated. Data comparing sparseness at times t1 and t2 for a single model neuron and the entire model population are plotted in Figure 8, E and F, respectively. We observed a modest but significant increase in population sparseness (16%, $p = 0.031$, paired t test). These data demonstrate that selectivities of individual neurons increased over time, with a time course consistent with the development of stSUP. Using a model, we could infer that population sparseness also showed a corresponding, but weaker, increase over time.

The effect of mixing different response types: a simulation analysis

It must be emphasized that the above recordings were exclusively based on well isolated SUs, separated into functional response types based on FRA shape. Mixing these neuronal populations for analyses typically washed out the inhibitory effects that we reported above. This mixing of different response types by MU recordings may explain why earlier experiments, based on MU data, did not observe the strong on-BF stSUP at loud levels or the large proportion of nonmonotonically intensity-tuned neurons as we did in the present study. It has been documented, for example, that the number of nonmonotonic units observed was a strong function of whether single- or multiunit spikes were analyzed (Sutter and Schreiner, 1995). To understand the effect of the relative proportions of O and V units on observing on-BF stSUP or level tuning as a function of recording quality, we simulated recording FRAs from a population of units having a mix of 64% O and 36% V units. Strong tonotopic organization was evident in our experimental data, but units were intermixed both in terms of BL and functional type, on the cortical surface as well as in depth in all recorded hemispheres. BLs and functional types of recorded units overlaid on the tonotopic map for one example hemisphere is shown in Figure 9. Extracellular recordings have a spatial resolution of $\sim 100 \mu\text{m}$; therefore, organization of sound level at a finer scale cannot be ruled out from our data. However, since the simulation also models extracellular recordings, an assumption of neurons not organized by BL or functional type is valid. Therefore, the model was based on tonotopically organized neurons that were not spatially segregated by BL or functional type (see Materials and Methods).

For simulated single units, there was a large proportion of O-shaped FRAs, reflecting actual data from awake marmoset A1 (Fig. 10A, SU). When we simulated a multiunit recording, with five SUs contributing to each MU response (MU-5), the FRA became progressively more I/V shaped (Fig. 10A, MU-5). For an MU recording with the summed contribution of 10 SUs (MU-10), the FRA was characteristically V shaped (Fig. 10A, MU-10). We quantified this resultant shape change by measuring monotonicity, defined using an MI (see Materials and Methods). Figure 10B shows the distributions of MIs when 200 recording sites of different qualities were simulated (SU, MU-5, and MU-10). This distribution gradually shifted from highly nonmonotonic for SUs to mostly monotonic for MUs. When we plotted the median MI of these distributions and the proportion of nonmonotonic units ($\text{MI} \leq 0.5$) in the simulated recordings (Fig. 10C), we also observed a systematic shift toward monotonicity with decreasing recording quality. The percentages of nonmonotonic units we observed in our simulations ($\sim 75\%$ for SUs, $\sim 25\%$ for MU-10s) were similar to the proportions observed in

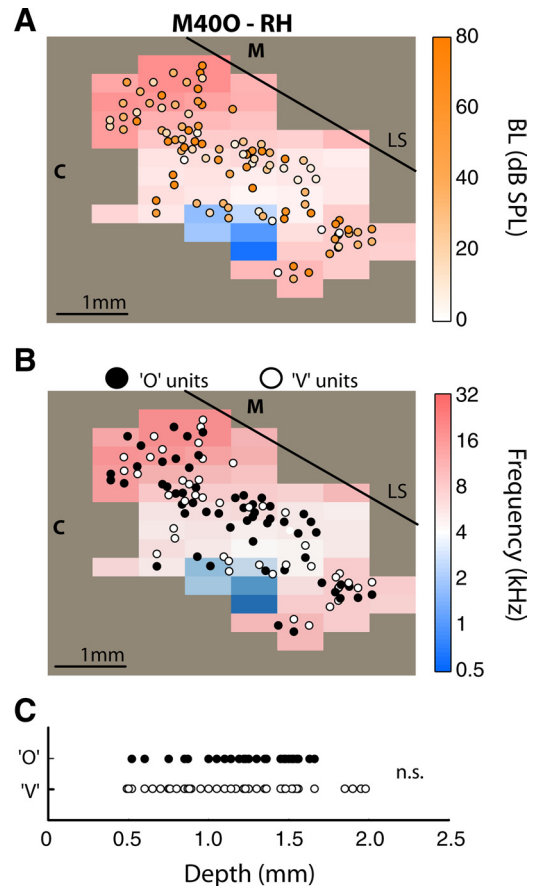


Figure 9. Topographical layout of best level and FRA shape in A1. **A**, BLs of neurons encountered in recordings from the right hemisphere of one marmoset monkey (M400). Colored discs are locations of units with well defined FRAs, and color intensity corresponds to the BL of the unit. Background color map corresponds to the tonotopic map for this hemisphere derived from both units shown in discs, and other units for which only frequency-tuning curves were measured (data not shown). **B**, Same as **A**, but units have been classified into O (black discs) or V (white discs) units based on FRA shape. No spatial organization based on BL or FRA shape was evident at this spatial scale. Black line corresponds to the location of the lateral sulcus (LS); M and C denote medial and caudal, respectively. **C**, O and V units were also intermixed across cortical depths, as measured by electrode depth relative to the surface. n.s., not significant.

previous experiments measured in A1 using single-unit recordings [78% in Pfingst and O'Connor (1981); 64% in Sadagopan and Wang (2008)] or multiunit recordings [20% in Sutter (2000); 23% in Phillips and Irvine (1981)].

An important factor causing this phenomenon was the physiological distinctness of the V unit population. We simulated MU recordings as the summed contributions of many SU responses. Although only $\sim 36\%$ of the underlying neurons had V-shaped FRAs, they contributed disproportionately to the summed activity because they exhibited about twice the peak firing rate of O units (median maximal firing rate for BF/BL tones: V unit, 42 spikes/s; O unit, 22 spikes/s). Additionally, any observable stSUP at loud sound levels was washed out by V unit contamination because V units had higher response rates at louder levels. Thus, it is crucial that units be well isolated during recording sessions and analyzed by separate response types to reveal the phenomena that we described in this study.

Discussion

In this study, we have described the contribution of iRFs to shaping the already finely tuned eRFs observed in A1s of awake ani-

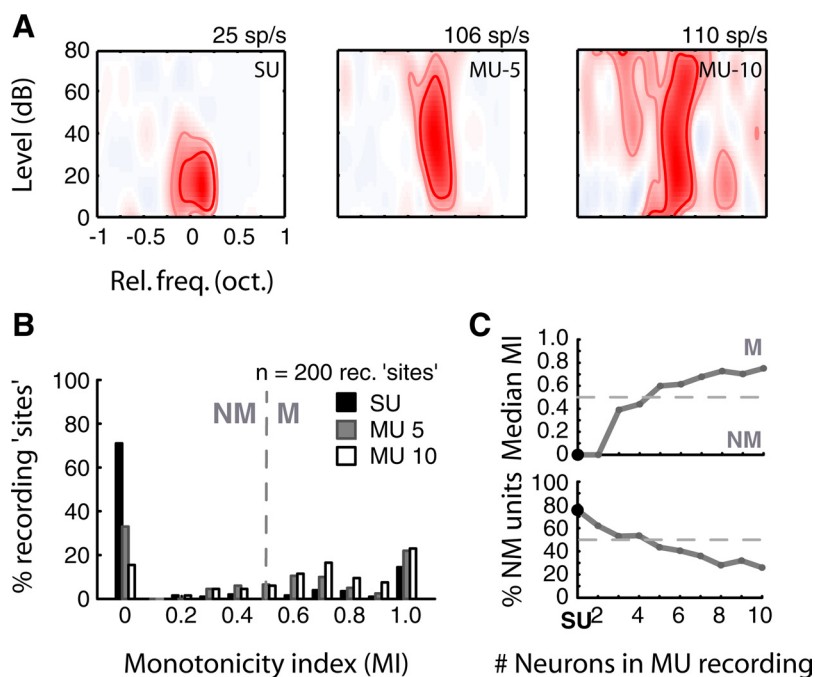


Figure 10. Simulation of the effect of O/V unit mix on observability of nonmonotonicity and stSUP. **A**, Examples of FRAs of SUs, simulated multiunits with the summed contributions of 5 single units (MU-5), and simulated multiunits with the summed contributions of 10 single units (MU-10). Color map corresponds to response strength contours as described earlier. Numbers are maximum response rates. **B**, Distribution of MIs derived from 200 recordings of SUs (black histogram), MU-5s (gray histogram), and MU-10s (white histogram). Gradual population shift toward monotonicity with decreasing signal quality is evident. **C**, Median MI of simulations from $n = 200$ recording sites (M, monotonic; NM, nonmonotonic) and proportion of nonmonotonic units observed in the population as a function of recording quality. When more than 5 SUs contribute to a given MU, we observed that a majority of the simulated recording sites exhibited monotonic rate-level functions.

mals. In an earlier study, we reported that sustained activation of A1 neurons could be obtained by using a neuron's preferred stimulus, tuned in several dimensions such as frequency, sound level, and amplitude and frequency modulations (Wang et al., 2005). The present study demonstrates a similar effect using pure-tone stimuli, with the additional observation that nonpreferred stimulus parameters in some dimensions could cause suppression. Had we used more well tuned complex stimuli, it is likely that we could have driven neurons to higher firing rates, but at the cost of not being able to precisely define eRFs and iRFs. The receptive fields of neurons are thus not completely defined by the axes explored in the present study. Therefore, while both simple and complex stimuli may be represented sparsely by the population of A1 neurons, firing rates of individual neurons in response to preferred stimuli are likely to be further modulated by other stimulus parameters. For this reason, a sparse representation does not conflict with sustained, strong responses reported in A1 (Wang et al., 2005). Rather, while only a fraction of neurons in the population are active for any particular stimulus (sparseness), the activation of these neurons is strong and sustained if the stimulus is well tuned to the neurons' eRF. A recent commentary addresses this issue in a more detailed manner (Willmore and King, 2009).

Relationship between stSUP and lateral inhibition

Three possible physiological mechanisms could underlie the observed stSUP: (1) reduction of excitatory drive to cortical neurons (e.g., through synaptic depression); (2) fast feed-forward inhibition from thalamus to cortex; and (3) lateral inhibition of cortical origin. For example, it is possible that feed-forward inhibition from the inferior colliculus or inhibition of intrinsic tha-

lamic origin (e.g., Suga et al., 1997) decreases the activity of MGB neurons, in turn reducing the excitatory drive to cortical neurons. In thalamocortical slices, EPSPs elicited by minimal stimulation at high frequencies show a rapid decrease in amplitude, perhaps as a result of depression at the thalamocortical synapse, but cause long-lasting depolarization of cortical neurons (Rose and Metherate, 2005). In awake animals, however, the excitability of A1 neurons is determined by a combination of feed-forward and lateral network activity. Therefore, it is unclear to what extent a reduction in feed-forward excitation, as a result of either a reduction of thalamic spiking activity or synaptic depression at the thalamocortical synapse, could cause cortical activity to fall significantly below spontaneous activity levels. A second possibility is feed-forward inhibition, whereby cortical inhibitory interneurons receive fast and direct inputs from MGB neurons that act to suppress responses to nonpreferred stimuli. Such an input might indeed contribute to the first stage of a two-stage inhibitory process (Sutter and Loftus, 2003) that may underlie the early level tuning seen in O neurons (Fig. 6).

Previous studies in visual cortex have suggested that single-stimulus suppression likely arises due to inhibitory inputs

(e.g., Monier et al., 2003; Martinez et al., 2002), suggesting that the observed stSUP is a consequence of active inhibition. The time course of the development of the observed inhibition may also suggest a cortical origin (Arthur et al., 1971; Ojima and Murakami, 2002). In slices, layer 2/3 pyramidal cells exhibit inhibition (evoked by stimulation of a connected fast-spiking interneuron) that develops to a steady-state value and is long lasting, consistent with the time scales of suppression observed in this study (Oswald et al., 2009). A recent study in the auditory cortex of mice used suppression of responses below the spontaneous rate as a measure of cortically derived inhibition, and changes in this suppression as a measure of inhibitory plasticity (Galindo-Leon et al., 2009). Therefore, taken together with the intracellular and pharmacological evidence described earlier (see Introduction), a combination of fast feed-forward and slower intracortical active inhibition, causing activity to fall below the spontaneous rate, is the more likely explanation in our case. Further pharmacological and intracellular data may be necessary to conclusively distinguish between these possibilities.

Contribution of inhibition in generating O units

We made the following relevant observations in A1 neurons of awake marmosets: (1) O unit FRAs were already O shaped in the earliest response bins (first 20 ms after onset latency), but exhibited higher level tuning widths; (2) on-BF late inhibition at loud levels, starting >30 ms after stimulus onset, further sharpened level tuning; and (3) inhibition at loud sound levels persisted for up to 100 ms after stimulus offset. Previously, a large number of nonmonotonic neurons was reported in the auditory thalamus of anesthetized cats by Rouiller et al. (1983) using a measure based

on slope changes of rate-level functions. In awake marmosets, a high proportion of neurons exhibiting nonmonotonic rate-level functions was observed in the medial geniculate body (E. Bartlett and X. Wang, unpublished data). Together, these data support the hypothesis that cortical O units receive thalamic inputs that are already broadly level tuned. The late inhibition we observed was similar to intracellular observations by Ojima and Murakami (2002) in anesthetized cats and may explain results from a pharmacological manipulation experiment that implicates cortical inhibitory sources in sound-level tuning (Wang et al., 2002). This late cortical inhibition may also be amenable to modification by training and may underlie the plasticity of level tuning that has been observed (Polley et al., 2006; Galindo-Leon et al., 2009). Therefore, our data from A1s of awake marmosets offer strong support for a two-stage inhibition process proposed earlier based on data derived using forward-masking paradigms in anesthetized cats (Calford and Semple, 1995; Sutter and Loftus, 2003), suggesting a locus in A1 for the second inhibitory stage.

Implications for stimulus design in studying A1 responses

The fact that iRFs are more broadly tuned than eRFs with effects lasting a longer duration should be carefully considered while designing stimuli to probe feed-forward processing in auditory cortex. Because of these wide and long-lasting inhibitory effects, dense stimuli are unlikely to drive auditory cortical neurons well (Blake and Merzenich, 2002). In addition, stimuli with rapid frequency contours are also suboptimal, as side-band suppression extending over time (Fig. 1*B,E*) imposes a strong low-pass characteristic on contour velocities. Similarly, loud sounds are likely to create long-lasting inhibitory effects that may prevent responses to other moderate-level stimuli for hundreds of milliseconds. Therefore, suboptimal stimuli may evoke undesired inhibitory effects that reduce responses to subsequent optimal stimuli as well. In the context of feed-forward processing, if one is interested in studying a processing stage subsequent to the tone-tuning stage, these concerns are amplified. The present characterization of the suppressive and inhibitory effects on A1 neurons may help in the design of stimulus sets to probe higher cortical function.

In our analysis, we have not considered the effects of spatial position on neural responses. Two broad hypotheses for the representation of sound location have been explored. First, individual neurons in auditory cortex may be tuned to restricted regions of acoustic space (e.g., Brugge et al., 1996; Schnupp et al., 2001). In this case, spatial location may be combinatorially represented in a manner similar to sound level: individual neurons have a preferred spatial location that elicits sustained spiking responses and nonpreferred spatial locations that may elicit weaker responses or suppression. Neurons may not be topographically organized along this parameter, and interactions between location and other parameters of the sound may be minimal. Alternatively, strong evidence for distributed codes based on spike timing or firing rates of neural populations is also available (Stecker and Middlebrooks, 2003; Stecker et al., 2005; Miller and Recanzone, 2009). A recent intracellular study found that individual neurons in auditory cortex receive panoramic inputs that are sharpened by spike threshold to result in tuned spatial receptive fields, with preferred locations eliciting earlier spikes (Chadderton et al., 2009). In this case, it is less clear how excitation and inhibition from the location axis would interact with the other axes described in this study. One possibility is that “preferred” or “nonpreferred” stimulus types may further enhance the differences in spike latency

observed between preferred and nonpreferred spatial locations. Further experiments are crucial to test how spatial location may be corepresented in A1 in our preparation.

In conclusion, in addition to the sharp tuning of the eRF, suppression by iRFs in the vicinity of a given neuron’s eRF confers high selectivity on a given neuron’s response in A1s of awake marmosets. As a result, responses of populations of A1 neurons are sparse (Hromádka et al., 2008), in the sense that only a few neurons respond for a given stimulus presentation. However, it must be noted that the responses of individual neurons are strong and robust when optimal stimuli that maximally stimulate eRF and minimally stimulate iRF are presented. These results have significant implications for future studies that are aimed at probing feed-forward response properties at stages that are higher than A1, as care needs to be taken in designing stimuli that could first pass through this highly restrictive A1 filter. It was suggested earlier that the effective region of stimulus space sampled by individual A1 neurons might shrink as response duration progresses (Wang, 2007). Here, we further suggest that, in addition to a narrowing of excitation, large regions of stimulus space also become inhibitory or suppressive, and contribute to a highly selective and sparse representation of sounds in A1s of awake animals (Fig. 7*D*).

References

- Abeles M, Goldstein MH Jr (1972) Responses of single units in the primary auditory cortex of the cat to tones and tone-pairs. *Brain Res* 42:337–352.
- Agamaite JA (1997) A quantitative characterization of the vocal repertoire of the common marmoset. Master’s thesis, Johns Hopkins University.
- Arthur RM, Pfeiffer RR, Suga N (1971) Properties of “two-tone inhibition” in primary auditory neurones. *J Physiol* 212:593–609.
- Barbour DL, Wang X (2003) Auditory cortical responses elicited in awake primates by random spectral stimuli. *J Neurosci* 23:7194–7206.
- Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature* 451:197–201.
- Blake DT, Merzenich MM (2002) Changes of A1 receptive fields with sound density. *J Neurophysiol* 88:3409–3420.
- Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. *J Neurophysiol* 36:1138–1158.
- Brugge JF, Reale RA, Hind JE (1996) The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *J Neurosci* 16:4420–4437.
- Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. *J Neurophysiol* 73:1876–1891.
- Chadderton P, Agapiou JP, McAlpine D, Margrie TW (2009) The synaptic representation of sound source location in auditory cortex. *J Neurosci* 29:14127–14135.
- Chen QC, Jen PH (2000) Bicuculline application affects discharge patterns, rate-intensity functions and frequency tuning characteristics of bat auditory cortical neurons. *Hear Res* 150:161–174.
- DiMattina C, Wang X (2006) Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 95:1244–1262.
- Evans EF, Whitfield IC (1964) Classification of unit responses in the auditory cortex of the unanesthetized and unrestrained cat. *J Physiol* 171:476–493.
- Firzlaff U, Schuller G (2001) Motion processing in the auditory cortex of the rufous horseshoe bat: role of GABAergic inhibition. *Eur J Neurosci* 14:1687–1701.
- Galindo-Leon EE, Lin FG, Liu RC (2009) Inhibitory plasticity in a lateral band improves cortical detection of natural vocalizations. *Neuron* 62:705–716.
- Hromádka T, Deweese MR, Zador AM (2008) Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol* 6:e16.
- Hubel DH, Henson CO, Rupert A, Galambos R (1959) Attention units in the auditory cortex. *Science* 129:1279–1280.
- Jen PH, Chen QC, Wu FJ (2002) Interaction between excitation and inhibi-

- tion affects frequency tuning curve, response size and latency of neurons in the auditory cortex of the big brown bat, *Eptesicus fuscus*. *Hear Res* 174:281–289.
- Lehky SR, Sejnowski TJ, Desimone R (2005) Selectivity and sparseness in the responses of striate complex cells. *Vision Res* 45:57–73.
- Loftus WC, Sutter ML (2001) Spectrotemporal organization of excitatory and inhibitory receptive fields of cat posterior auditory field neurons. *J Neurophysiol* 86:475–491.
- Lu T, Liang L, Wang X (2001) Neural representation of temporally asymmetric stimuli in the auditory cortex of awake primates. *J Neurophysiol* 85:2364–2380.
- Martinez LM, Alonso JM, Reid RC, Hirsch JA (2002) Laminar processing of stimulus orientation in cat visual cortex. *J Physiol* 540:321–333.
- Miller LM, Recanzone GH (2009) Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proc Natl Acad Sci U S A* 106:5931–5935.
- Monier C, Chavane F, Baudot P, Graham LJ, Frégnac Y (2003) Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron* 37:663–680.
- O'Connor KN, Petkov CI, Sutter ML (2005) Adaptive stimulus optimization for auditory cortical neurons. *J Neurophysiol* 94:4051–4067.
- Ojima H, Murakami K (2002) Intracellular characterization of suppressive responses in supragranular pyramidal neurons of cat primary auditory cortex in vivo. *Cereb Cortex* 12:1079–1091.
- Oswald AM, Doiron B, Rinzal J, Reyes AD (2009) Spatial profile and differential recruitment of GABA-B modulate oscillatory activity in auditory cortex. *J Neurosci* 29:10321–10334.
- Pelleg-Toiba R, Wollberg Z (1989) Tuning properties of auditory cortex cells in the awake squirrel monkey. *Exp Brain Res* 74:353–364.
- Pfingst BE, O'Connor TA (1981) Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J Neurophysiol* 45:16–34.
- Phillips DP, Irvine DR (1981) Responses of single neurons in physiologically defined primary auditory cortex (A1) of the cat: frequency tuning and responses to intensity. *J Neurophysiol* 45:48–58.
- Polley DB, Steinberg EE, Merzenich MM (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci* 26:4970–4982.
- Razak KA, Fuzessery ZM (2009) GABA shapes selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex. *J Neurophysiol* 102:1366–1378.
- Rose HJ, Metherate R (2005) Auditory thalamocortical transmission is reliable and temporally precise. *J Neurophysiol* 94:2019–2030.
- Rouiller E, de Ribaupierre Y, Morel A, de Ribaupierre F (1983) Intensity functions of single unit responses to tone in the medial geniculate body of cat. *Hear Res* 11:235–247.
- Sadagopan S, Wang X (2008) Level invariant representation of sounds by populations of neurons in primary auditory cortex. *J Neurosci* 28:3415–3426.
- Sadagopan S, Wang X (2009) Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. *J Neurosci* 29:11192–11202.
- Schnupp JW, Msrsc-Flogel TD, King AJ (2001) Linear processing of spatial cues in primary auditory cortex. *Nature* 414:200–204.
- Scholl B, Gao X, Wehr M (2008) Level dependence of contextual modulation in auditory cortex. *J Neurophysiol* 99:1616–1627.
- Shamma SA, Symmes D (1985) Patterns of inhibition in auditory cortical cells in awake squirrel monkeys. *Hear Res* 19:1–13.
- Stecker GC, Middlebrooks JC (2003) Distributed coding of sound locations in the auditory cortex. *Biol Cybern* 89:341–349.
- Stecker GC, Harrington IA, Middlebrooks JC (2005) Location coding by opponent neural populations in auditory cortex. *PLoS Biol* 3:e78.
- Suga N (1965) Functional properties of auditory neurones in the cortex of echo-locating bats. *J Physiol* 181:671–700.
- Suga N, Zhang Y, Yan J (1997) Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *J Neurophysiol* 77:2098–2114.
- Sutter ML (2000) Shapes and level tolerances of frequency tuning curves in primary auditory cortex: quantitative measures and population codes. *J Neurophysiol* 84:1012–1025.
- Sutter ML, Loftus WC (2003) Excitatory and inhibitory intensity tuning in auditory cortex: evidence for multiple inhibitory mechanisms. *J Neurophysiol* 90:2629–2647.
- Sutter ML, Schreiner CE (1995) Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multi-neuron recordings. *J Neurophysiol* 73:190–204.
- Sutter ML, Schreiner CE, McLean M, O'Connor KN, Loftus WC (1999) Organization of inhibitory frequency receptive fields in cat primary auditory cortex. *J Neurophysiol* 82:2358–2371.
- Tan AY, Atencio CA, Polley DB, Merzenich MM, Schreiner CE (2007) Unbalanced synaptic inhibition can create intensity-tuned auditory cortex neurons. *Neuroscience* 146:449–462.
- Wang J, McFadden SL, Caspary D, Salvi R (2002) Gamma-aminobutyric acid circuits shape response properties of auditory cortex neurons. *Brain Res* 944:219–231.
- Wang X (2007) Neural coding strategies in auditory cortex. *Hear Res* 229:81–93.
- Wang X, Kadia SC (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86:2616–2620.
- Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435:341–346.
- Willmore BD, King AJ (2009) Auditory cortex: representation through sparsification? *Curr Biol* 19:R1123–R1125.
- Wu GK, Li P, Tao HW, Zhang LI (2006) Nonmonotonic synaptic excitation and imbalanced inhibition underlying cortical intensity tuning. *Neuron* 52:705–715.
- Zhang LI, Tan AY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424:201–205.