Systems/Circuits

Somatostatin-Expressing Interneurons in the Auditory Cortex Mediate Sustained Suppression by Spectral Surround

Anna A. Lakunina, Matthew B. Nardoci, Yashar Ahmadian, and Santiago Jaramillo
Institute of Neuroscience and Department of Biology, University of Oregon, Eugene, OR 97403

Sensory systems integrate multiple stimulus features to generate coherent percepts. Spectral surround suppression, the phenomenon by which sound-evoked responses of auditory neurons are suppressed by stimuli outside their receptive field, is an example of this integration taking place in the auditory system. While this form of global integration is commonly observed in auditory cortical neurons, and potentially used by the nervous system to separate signals from noise, the mechanisms that underlie this suppression of activity are not well understood. We evaluated the contributions to spectral surround suppression of the two most common inhibitory cell types in the cortex, parvalbumin-expressing (PV+) and somatostatin-expressing (SOM$^+$) interneurons, in mice of both sexes. We found that inactivating SOM$^+$ cells, but not PV+ cells, significantly reduces sustained spectral surround suppression in excitatory cells, indicating a dominant causal role for SOM$^+$ cells in the integration of information across multiple frequencies. The similarity of these results to those from other sensory cortices provides evidence of common mechanisms across the cerebral cortex for generating global percepts from separate features.

Key words: sensory processing; auditory cortex; cortical inhibition

---

Significance Statement

To generate coherent percepts, sensory systems integrate simultaneously occurring features of a stimulus, yet the mechanisms by which this integration occurs are not fully understood. Our results show that neurochemically distinct neuronal subtypes in the primary auditory cortex have different contributions to the integration of different frequency components of an acoustic stimulus. Together with findings from other sensory cortices, our results provide evidence of a common mechanism for cortical computations used for global integration of stimulus features.

Introduction

We experience the acoustic world as a set of coherent auditory objects, although sensory organs in our ears detect the individual frequency components of sound. This implies the existence of mechanisms in the auditory system for integrating information across frequency channels to create global percepts (Metherate et al., 2005). One example of global integration in the auditory system is spectral surround suppression of sound-evoked responses, the phenomenon by which a neuron's sound-driven responses are suppressed by the presence of sound frequencies outside its classical receptive field. The activity of neurons in the auditory cortex of mammals is commonly suppressed by pure tones (PTs)

outside the neurons' receptive fields (Abeles and Goldstein, 1972; Nelken et al., 1994), or sounds of high bandwidth (Rauschecker and Tian, 2004; O'Connor et al., 2010; H. Li et al., 2019), yet the mechanism by which the auditory system integrates information across many frequency channels to produce these suppressed responses in the cortex is not well understood. Specifically, it is not clear to what extent this suppression is inherited from upstream auditory areas or mediated by cortical circuits.

Here, we investigated the role of inhibitory auditory cortical circuits in mediating spectral surround suppression. When classified according to gene-expression patterns, distinct classes of inhibitory interneurons within a particular cortical region display different physiological responses and anatomic connectivity (Kvitsiani et al., 2013; Moore and Wehr, 2013; L. Y. Li et al., 2014, 2015; Kuchibhotla et al., 2017; Liang et al., 2019), suggesting a difference in function between these cell types (Adesnik et al., 2012; Natan et al., 2015; Kato et al., 2017; Phillips et al., 2017). To evaluate the contributions of distinct sources of inhibition to cortically mediated surround suppression of evoked responses to high bandwidth sound stimuli, we studied the two most common genetically defined types of cortical inhibitory interneurons, parvalbumin-expressing (PV$^+$) and somatostatin-expressing (SOM$^+$) cells (Rudy et al., 2011). We found that SOM$^+$ cells play a more dominant role in mediating sustained

auditory cortical spectral surround suppression than PV$^+$ cells, suggesting a mechanistic role of SOM$^+$ cells in the global integration of auditory stimuli.

Materials and Methods

Animals
A total of 31 transgenic adult mice of both sexes (4 SOM-Cre, 7 PV-Cre, 16 SOM::ChR2, and 4 PV::ChR2), age 3-7 months, were used in this study. The PV-Cre and SOM-Cre driver lines (008069 and 013044 from The Jackson Laboratory) were used in combination with viral transfection, or crossed with LSL-ChR2 mice (012569 from The Jackson Laboratory) to produce mice expressing channelrhodopsin-2 (ChR2) in PV$^+$ or SOM$^+$ inhibitory interneurons. All procedures were conducted in accordance with National Institutes of Health standards and were approved by the University of Oregon Institutional Animal Care and Use Committee.

Experimental details
Auditory stimuli
Electrophysiology was conducted inside single-walled sound-isolation boxes (IAC Acoustics). All auditory stimuli were generated using Python software developed in-house (https://taskontrol.readthedocs.io/) and were delivered from a free-field speaker (MF1 Multi-Field Magnetic Speakers, Tucker-Davis Technologies) facing the mouse's ear contralateral to the recording. Mice were head-fixed on top of a freely-moving wheel, leaving them free to move and run while their heads remained stationary.

Recorded cells were tested for their characteristic frequency and peak response to amplitude modulation (AM). Frequency tuning was determined through the presentation of 100 ms PTs at 16 frequencies (ranging from 2 to 40 kHz) at 55 dB SPL. This sound intensity was selected to match the power of the individual frequency components in 70 dB SPL white noise (WN). The preferred frequency was selected by fitting a Gaussian curve to the cell's frequency response curve and selecting the peak as the estimated best frequency. AM responses were determined through presentation of 500 ms of sinusoidally amplitude-modulated WN with a modulation depth of 100% at 5 modulation rates (ranging from 4 to 64 Hz). AM sounds were presented at a peak intensity of 70 dB SPL. The best modulation rate was then selected as the one that gave the greatest sustained firing rate for each cell.

To determine the degree of surround suppression for each cell, we recorded responses to bandpass-filtered WN centered at the cell's preferred frequency and amplitude modulated at the cell's preferred rate. Stimuli were presented at 7 different bandwidths: 0 pure tone (PT), 0.25, 0.5, 1, 2, 4 octaves, and WN. The WN stimulus had an intensity of 70 dB SPL. All other stimuli were bandpass-filtered versions of this WN stimulus, with the exception of the PT stimulus, which was generated separately as a pure sinusoidal wave. This difference in sound generation made it difficult to precisely match the amplitude of the PT stimulus to that of the other stimuli, potentially affecting the interpretability of our results. For this reason, we focused our analysis of surround suppression on the responses to bandpass-filtered noise. To test for intensity tuning, additional trials were performed with the WN stimulus at 50 dB SPL. All sound stimuli were 1000 ms long. Cells were selected for analysis of surround suppression if they were frequency tuned ($R^2 > 0.1$ of Gaussian fit) and the estimated preferred frequency was within 0.3 octaves of the center frequency used during the bandpass-filtered WN stimuli.

Surgical implant for head-fixed recordings
Animals were anesthetized with isoflurane through a nose-cone on a stereotaxic apparatus. Mice were surgically implanted with a head-bar to allow for head-fixed recordings. Bilateral craniotomies and durotomies were also performed over the auditory cortex to allow for acute recordings. Craniotomies were centered at 2.8 mm posterior to bregma and 4.4 mm from midline. All animals were monitored and recovered fully before electrophysiological experiments.

Viral injections
AAV5.CBA.Flex.ArchT-tdtomato.WPRE.SV40 (Addgene viral prep #28 305-AAV5) (Han et al., 2011) was injected bilaterally into the auditory cortex of PV-Cre and SOM-Cre mice. Two 90 nL injections were done per hemisphere at 2.6 and 3.0 mm posterior to bregma, 4.4 mm from midline, and at a depth of 1.0 mm from pia. The needle was kept at the injection depth for 3 min following injection. The virus was given at least 2 weeks to express before electrophysiological recordings.

Neural recordings
Electrical signals were collected using an RHD2000 acquisition system (Intan Technologies) and OpenEphys software (www.open-ephys.org), using silicon probe electrodes (A4x2-tet configuration from NeuroNexus). Multiple penetrations were made spanning the mediolateral axis of the craniotomy. Probes were covered in fluorescent dye (DiO, DiI, or DiD; Vybrant cell-labeling solution) before each penetration to verify placement at the conclusion of the experiment. During the experiment, the mouse was head-fixed inside the rig and the protective covering over the craniotomy was removed, exposing the brain. The probe was moved into place, such that its shanks were parallel to the midline of the skull. The probe was held in a vertical position and lowered vertically into the cortex until spikes were detected. We recorded at multiple depths on each penetration, with recording sites typically 100–150 μm apart to avoid recording from the same cells twice.

Photoidentification of PV$^+$ and SOM$^+$ neurons
PV$^+$ and SOM$^+$ cells were identified by evaluating neural responses to the first 10 ms of light pulses in PV::ChR2 and SOM::ChR2 mice. Light pulses were 100 ms long with interstimulus interval of 900 ms, but only the onset responses were used for photoidentification to rule out effects from synaptic transmission. Blue light (wavelength 470 nm) was delivered via an optical fiber attached to the silicon probe electrodes with the tip ~200 μm from the topmost recording sites (fiber diameter 50 μm). Light power was 1.5-2.5 mW at the fiber tip. PV$^+$ and SOM$^+$ cells were identified by their significant ($p < 0.001$) increase in firing rate immediately following light presentation. Putative excitatory cells were identified as cells in SOM::ChR2 mice that did not have positive responses to the light pulses and whose spike widths, defined as the time difference between the sodium peak ("peak") and the potassium peak ("trough"), were >0.4 ms. Because the spike shapes of excitatory cells and SOM$^+$ cells can be very similar, putative excitatory cells were selected only from SOM::ChR2 mice to exclude SOM$^+$ cells by their laser response and exclude PV$^+$ cells by spike shape.

Inactivation of PV$^+$ and SOM$^+$ neurons
PV$^+$ and SOM$^+$ cells were inactivated during sound presentation with 1300 ms light pulses (light onset was 100 ms before sound onset, light offset was 200 ms after sound offset). Green light (wavelength 520 nm) was delivered via an optical fiber attached to the silicon probe electrodes with the tip ~900 μm from the topmost recording sites (fiber diameter 200 μm). Light power was 5 mW at the fiber tip. Laser was presented for 50% of trials. Laser and nonlaser trials were randomly interleaved. Laser-induced changes in baseline firing rate were calculated using the first 50 ms after laser onset.

Histology
At the conclusion of the experiment, animals were deeply anesthetized with euthasol and perfused through the heart with 4% PFA. Brains were extracted and left in 4% PFA for at least 24 h before slicing. Brains were sliced under PBS on a vibratome with a slice thickness of 50 μm. Brain slices were imaged with a fluorescent microscope (Axio Imager 2, Carl Zeiss) with a 2.5× objective. To determine the locations of our recordings, we manually registered each histology slice containing dye fluorescence to the corresponding coronal section in the Allen Mouse Common Coordinate Framework (Common Coordinate Framework version 3, 2015 Allen Institute for Brain Science, Allen Brain Atlas API; http://brain-map.org/api/index.html). Recordings were considered for analysis only if they were localized to auditory cortical areas.

Immunohistochemistry
Animals were deeply anesthetized with Euthasol and perfused through the heart with 4% PFA, and the brains were postfixed overnight and cryoprotected in 30% sucrose. Sections 30 μm thick were blocked in 10% donkey serum in PBS for 1 h. Sections were then incubated for 24 h in mouse anti-PV (1:4000, Millipore MAB1572) or rat anti-SOM (1:50,
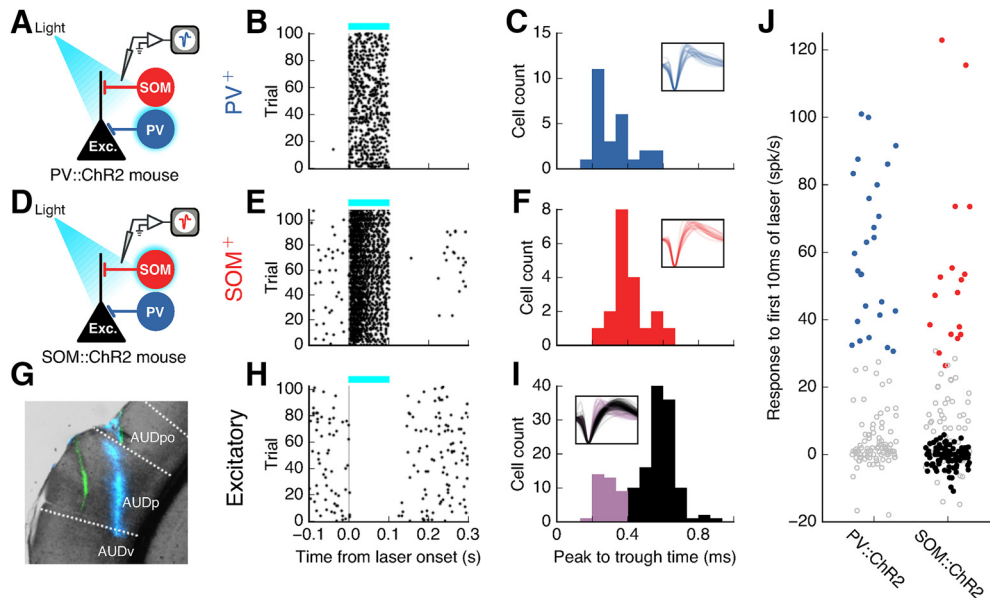
Figure 1. Photoidentification of auditory cortical cell types. ***A***, Identification of PV$^+$ cells in PV::ChR2 mice by their responses to laser stimulation. ***B***, Response of an example identified PV$^+$ cell to 100 ms pulses of blue laser (blue bar). ***C***, Spike widths of all identified PV$^+$ cells ($N = 26$). Inset, Normalized spike waveforms of all identified PV$^+$ cells. Voltage traces are 1.33 ms long. ***D***, Identification of SOM$^+$ cells in SOM::ChR2 mice by their responses to laser stimulation. ***E***, Response of an example identified SOM$^+$ cell to 100 ms pulses of blue laser (blue bar). ***F***, Spike widths of all identified SOM$^+$ cells ($N = 19$). ***G***, Coronal brain slice (AP: −3.18 mm) showing two electrode tracks: green (DiO) and blue (DiD). Recordings were performed at multiple sites along each penetration. AUDp, Primary auditory area; AUDpo, posterior auditory area; AUDv, ventral auditory area, according to Allen Mouse Brain Atlas. ***H***, Response of an example putative excitatory cell to 100 ms pulses of blue laser (blue bar). ***I***, Spike widths of all cells that did not have positive responses to laser from SOM::ChR2 animals. Black represents cells classified as putative excitatory cells ($N = 116$). Purple represents cells with narrow spikes (peak to trough < 0.4 ms, $N = 37$) which were not classified as putative excitatory cells. ***J***, Laser-evoked change in firing for all frequency-tuned cells ($N = 122$ from PV::ChR2 mice, 204 from SOM::ChR2 mice) during the first 10 ms of the laser pulse. Blue represents identified PV$^+$ cells. Red represents identified SOM$^+$ cells. Black represents putative excitatory cells. Gray represents cells not falling into any of the above three categories. PV$^+$ and SOM$^+$ cells were identified by their positive, low-latency responses to blue laser ($p < 0.001$, Wilcoxon rank-sum test). Two SOM$^+$ cells with laser-induced firing rates of >200 Hz are not shown.

Abcam M09204). The sections were then incubated for 2 h in donkey anti-mouse (1:800, Thermo Fisher Scientific/Invitrogen SA5-10 166) or donkey anti-rat (1:800, Thermo Fisher Scientific Invitrogen SA5-10 026). Brain slices were imaged with a fluorescent microscope (Axio Imager 2, Carl Zeiss) with a 10× objective. To quantify the specificity of expression, cells were identified in separate fluorescent channels and subsequently scored for colocalization.

Data analysis

Spike sorting
Spiking activity was detected by applying a low threshold (40–45 μV) to bandpass-filtered (300–6000 Hz) continuous data. Spiking activity of single units was isolated offline using the automated expectation maximization clustering algorithm Klustakwik (Kadir et al., 2014). Isolated clusters were only included in the analysis if $<2\%$ of interspike intervals were <2 ms. Clusters with 2%-5% of interspike intervals <2 ms were automatically refined by iteratively removing the spike with the largest Mahalanobis distance to the cluster centroid in feature space until the cluster had $<2\%$ of interspike intervals <2 ms. We also calculated a spike quality index, defined as the ratio between the peak amplitude of the waveform and the average variance, calculated using the channel with the largest amplitude. Cells were only included in the analysis if they had a spike quality index >2.5.

Suppression index (SI)
To determine the degree to which cells are suppressed by sounds of high bandwidth, we calculated an SI as follows:

$$\text{SI} = \frac{R_p - R_w}{R_p}$$

Here, R_p is the cell's peak firing rate and R_w is the cell's firing rate in response to WN, the highest bandwidth stimulus presented. This SI is similar to one previously used to quantify surround suppression in the

visual cortex (Adesnik et al., 2012; Ayaz et al., 2013). R_p and R_w were calculated using fits from a subtractive surround model.

Facilitation index
To determine the degree to which cell responses increase as the bandwidth of the sound stimulus increases from zero (PT), we calculated a facilitation index as follows:

$$\text{FI} = \frac{R_p - R_{pt}}{R_p}$$

Here, R_p is the cell's peak firing rate and R_{pt} is the cell's firing rate in response to the PT stimulus at the cell's preferred frequency. R_p and R_{pt} were calculated using fits from a subtractive surround model that included responses to PTs.

PT versus WN index
To determine the degree to which cells exhibited a preference for PTs (the lowest bandwidth stimulus presented) over WN (the highest bandwidth stimulus presented), we calculated a PT versus WN index as follows:

$$\frac{R_{pt} - R_w}{R_{pt} + R_w}$$

Here, R_{pt} is the cell's firing rate in response to the PT stimulus at the cell's preferred frequency, and R_w is the cell's firing rate in response to the WN stimulus.

Intensity index
To determine the degree to which cells were tuned to different intensities of sound, we calculated an intensity index as follows:
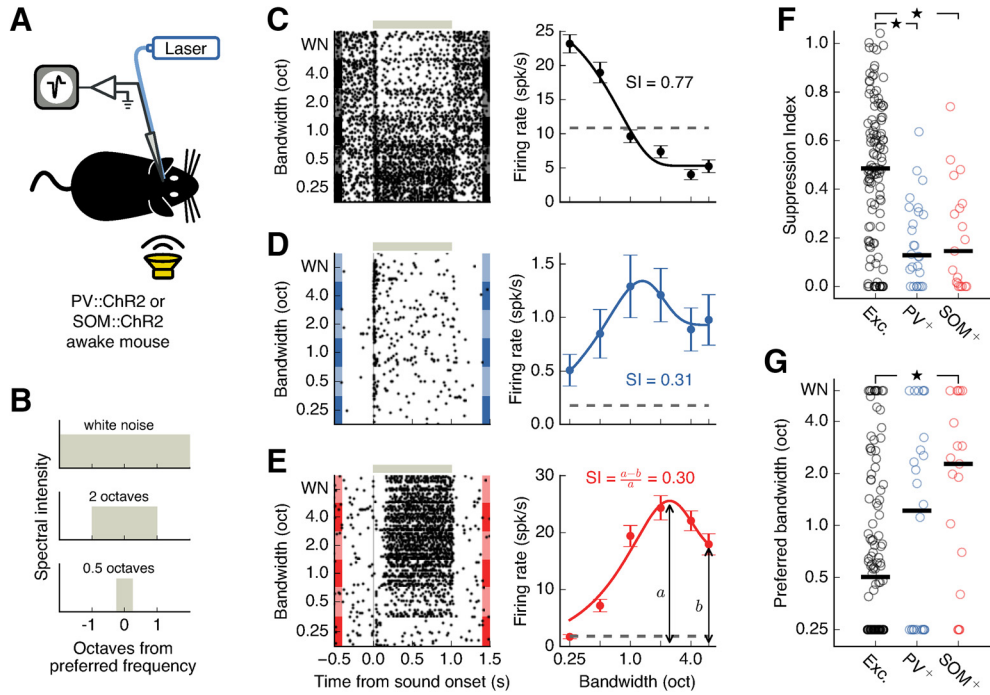
Figure 2. Inhibitory interneurons in auditory cortex exhibit less spectral surround suppression than excitatory neurons. **A**, Setup for electrophysiological recordings from the auditory cortex during contralateral sound presentation. **B**, Frequency spectra of auditory stimuli used to estimate bandwidth tuning. **C**, Left, Response of the example putative excitatory cell depicted in Figure 1H to amplitude-modulated sounds of different bandwidths. Gray bar represents the sound duration. Right, Bandwidth tuning curve of the sustained response (200-1000 ms after sound onset). Dashed line indicates spontaneous firing rate. Error bars indicate SEM. This cell shows a typical surround-suppressed response: lower firing for high bandwidth stimuli than for low bandwidth stimuli. **D**, Same as in **C** for the PV$^+$ cell depicted in Figure 1B. **E**, Same as in **C** for the SOM$^+$ cell depicted in Figure 1E. **F**, Spectral surround suppression for all recorded putative excitatory ($N = 116$), PV$^+$ ($N = 26$), and SOM$^+$ cells ($N = 19$). Black bars represent medians. Inhibitory cells were significantly less suppressed than excitatory cells (excitatory PV$^+$ $p = 5.9 \times 10^{-6}$, excitatory SOM$^+$ $p = 0.00039$, PV$^+$-SOM$^+$ $p = 0.863$, Wilcoxon rank-sum test). **G**, SOM$^+$ cells reach their peak firing rates at significantly higher bandwidths than excitatory cells (excitatory PV$^+$ $p = 0.327$, excitatory SOM$^+$ $p = 0.0083$, PV$^+$-SOM$^+$ $p = 0.260$, Wilcoxon rank-sum test). Black bars represent medians. *$p < 0.017$ (α value adjusted for multiple comparisons).

$$II = \frac{R_h - R_l}{R_h + R_l}$$

Here, R_h is the cell's firing rate in response to the high-intensity stimulus, and R_l is the cell's firing rate in response to the low-intensity stimulus. Intensity indices for each cell were calculated using the responses to the PT and WN stimuli to determine the extent to which surround suppression is explainable by intensity tuning.

Subtractive surround model fitting
Neuronal responses were modeled using a subtractive surround model similar to a divisive normalization model previously used to describe surround suppression in the visual cortex (Ayaz et al., 2013). In this model, two Gaussian fields provide input to each cell: a driving field G_d and a suppressive field G_s. The sum of their responses to sound stimuli with bandwidth b are as follows:

$$D(b) = \int_{-b}^{b} G_d(x)dx = \mathrm{erf}\left(\frac{b}{\sqrt{2}\sigma_d}\right)$$

$$S(b) = \int_{-b}^{b} G_s(x)dx = \mathrm{erf}\left(\frac{b}{\sqrt{2}\sigma_s}\right)$$

Here, σ_d and σ_s are the extents of the two Gaussian fields. The cell's response to sounds of different bandwidths is then computed as the difference of the two field responses:

$$R(b) = R_0 + R_d D(b) - R_s S(b)$$

Here, R_0 is the spontaneous firing rate, and R_d and R_s are the strengths of the driving and suppressive fields, respectively. The model was fit by minimizing the mean squared error between predicted and observed firing rates. A subtractive surround model was selected to model the responses of auditory cortical neurons because a divisive normalization model was not able to accurately capture the responses of cells that were suppressed below their spontaneous firing rate.

Mechanistic model
The feedforward neural model presented in Figure 9 consists of an output excitatory neuron that receives three types of inputs: (1) feedforward excitation from other cells (including thalamic inputs), (2) feedforward inhibition from PV$^+$ cells, and (3) feedforward inhibition from SOM$^+$ cells. The response of the output excitatory cell (R) is calculated as the weighted sum of the input $r(x)$ from each frequency channel x, passed through a rectifying nonlinearity as follows:

$$R = f\left(\sum_{\phi}\sum_{x} w_\phi(x)\, r(x)\right)$$

$$f(y) = \max(0, y)$$

where the weights $w_\phi(x)$ for each input type $\phi = \{Th, PV, SOM\}$ follow a Gaussian profile defined by two parameters: the maximum strength (A_ϕ) and the breadth of integration across frequency channels (σ_ϕ) as follows:

$$w_\phi(x) = A_\phi e^{-\frac{1}{2}\left(x/\sigma_\phi\right)^2}$$

The parameters σ_ϕ effectively combine the response-tuning width of presynaptic cells of type ϕ with the range of those cells' synaptic projections to excitatory cortical cells. The input $r(x)$ is simulated as a rectangular function of unit amplitude, centered around location $x = 0$. The width of this function defines the bandwidth of the stimulus. When thinking of $r(x)$ as representing thalamic inputs to the auditory cortex, our model neglects to account for any surround suppression in the thalamus (or more generally, in the excitatory input to excitatory cells).

We constructed a model with 101 frequency channels covering the entire tonotopic axis of the auditory cortex, which in mice covers ~6 octaves. For the examples in Figure 9B, C, the following synaptic strength parameters were used: $A_{PV} = -25$, $A_{SOM} = -25$, $A_{Th} = 100$, $\sigma_{PV} = 0.8$ octaves, $\sigma_{SOM} = 1.6$ octaves, and $\sigma_{Th} = 0.4$ octaves. These parameters were selected based on the results of Kato et al. (2017), which found that SOM$^+$ cells have a larger frequency tuning width and larger synaptic projection range than PV$^+$ cells. We chose to base our model on this study over others due to the similarity in experimental techniques, chiefly the use of long-duration sound stimuli in awake animals for measuring tuning widths. The additional simulations exploring the parameters space, shown in Figure 9,D, E used A_{PV} and A_{SOM} in the range 1%–30% of A_{Th}, σ_{Th} in the range 20%–100% of σ_{PV}, and σ_{SOM} in the range 100%–200% of σ_{PV}, with A_{Th} and σ_{PV} fixed at the values given above.

Statistics

To test for statistically significant effects across cells (e.g., comparing SIs across excitatory, PV$^+$, and SOM$^+$ cells), we used the nonparametric Wilcoxon rank-sum test because the data were rarely normally distributed and of equal variance. To test for significant effects within cells (e.g., testing effects of laser inactivation), we used the nonparametric Wilcoxon signed-rank test when testing across two conditions, or the Kruskal–Wallis test when testing across more than two conditions. To test for significant correlations between two continuous variables, we computed a least-squares linear regression. A Bonferroni correction was applied in situations where multiple comparisons were made.

Results

Inhibitory cells exhibit less spectral surround suppression than excitatory cells

To identify populations of cells involved in global integration, we first evaluated the degree to which distinct cell types exhibit surround suppression. To evaluate the responses of PV$^+$ and SOM$^+$ cells, we obtained single-unit extracellular recordings in the auditory cortex (Fig. 1G) of awake mice expressing the light-gated ion channel ChR2 in either PV$^+$ cells (Fig. 1A) or SOM$^+$ cells (Fig. 1D). We optogenetically identified PV$^+$ and SOM$^+$ cells as those exhibiting positive responses to the first 10 ms of pulses of blue light (Fig. 1J). Putative excitatory cells were defined as neurons that did not have a significant positive response to laser stimulation and had a spike width >0.4 ms (Fig. 1I,J).
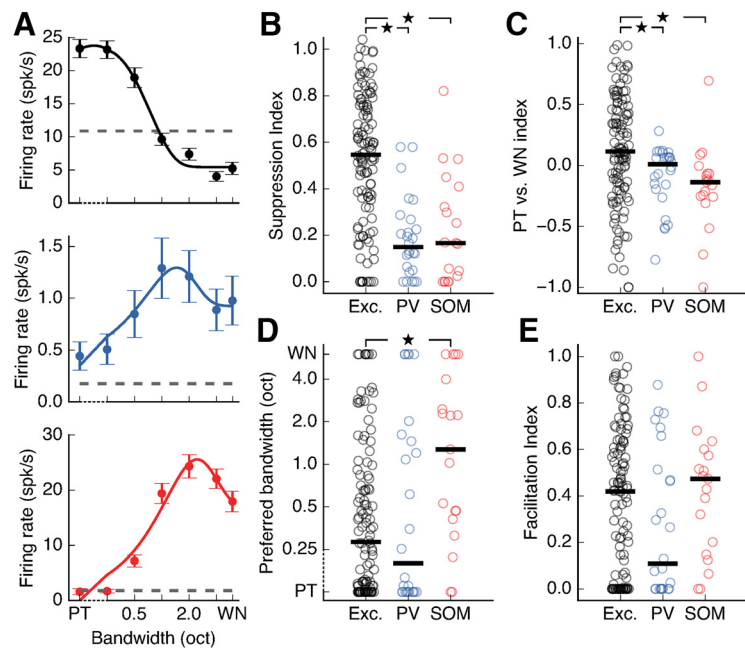

Figure 3. Inclusion of PT stimuli in analysis of suppression does not affect results. *A*, Bandwidth tuning curves for the cells depicted in Figure 2C-E when including responses to PTs (bandwidth = 0). Our estimation of suppression for these cells is minimally affected by the inclusion of these data points. Dashed axis indicates that data points from 0 (PT) to 0.25 octaves are plotted on a linear scale. *B*, Spectral surround suppression for all cells depicted in Figure 2F when responses to PTs are included in the calculation of SI. Inhibitory cells still show significantly less suppression than excitatory cells (excitatory PV$^+$ $p = 1.9 \times 10^{-7}$, excitatory SOM$^+$ $p = 5.3 \times 10^{-5}$, PV$^+$-SOM$^+$ $p = 0.918$, Wilcoxon rank-sum test). Black bars represent medians. *C*, When comparing responses to PTs and WN, excitatory cells show a significantly higher preference for PT stimuli compared with PV$^+$ or SOM$^+$ cells (excitatory PV$^+$ $p = 0.0088$, excitatory SOM$^+$ $p = 0.0026$, PV$^+$-SOM$^+$ $p = 0.113$, Wilcoxon rank-sum test). A positive index indicates that the cell had a larger response to PTs than WN. Black bars represent medians. *D*, The differences in preferred bandwidth seen in Figure 2G remain consistent when responses to PTs are taken into account (excitatory PV$^+$ $p = 0.920$, excitatory SOM$^+$ $p = 0.0044$, PV$^+$-SOM$^+$ $p = 0.056$, Wilcoxon rank-sum test). Dashed axis indicates that data points from 0 (PT) to 0.25 octaves are plotted on a linear scale. Black bars represent medians. *E*, There is no significant difference across cell types in facilitation of responses by increasing bandwidth (excitatory PV$^+$ $p = 0.187$, excitatory SOM$^+$ $p = 0.498$, PV$^+$-SOM$^+$ $p = 0.089$, Wilcoxon rank-sum test). Black bars indicate medians. *$p < 0.017$ (α value adjusted for multiple comparisons).

Because cells that do not respond to laser in PV::ChR2 mice could be SOM$^+$ or excitatory, and the shapes of these cell types largely overlap, we focused our initial analysis on putative excitatory cells from SOM::ChR2 mice, in which excitatory cells are easier to distinguish. To quantify the degree of spectral surround suppression in each cell, we presented sound stimuli of different bandwidths centered on each cell's preferred frequency to the ear contralateral to the recorded hemisphere (Fig. 2A,B). We then fit the cell's responses using a subtractive surround model (see Materials and Methods) and computed an SI defined as the normalized difference between the cell's maximum sound-evoked firing rate and its firing rate in response to white noise (WN), following previous studies exploring a similar effect in the visual system (Adesnik et al., 2012; Ayaz et al., 2013).

Figure 2C depicts the responses of a typical putative excitatory cell that displays strong spectral surround suppression. While the sustained responses of this cell to a low bandwidth stimulus at its preferred frequency were large, sound responses decreased with increasing bandwidth, although the power in the preferred frequency remained the same. We found that surround-suppressed responses were common in excitatory cells in the auditory cortex, with the population showing a median SI of 0.49 (Fig. 2F). SI was negatively correlated with cortical depth for excitatory cells, showing a median SI of 0.60 for supragranular layers and 0.43 for infragranular layers ($r = -0.259$, $p = 0.006$,
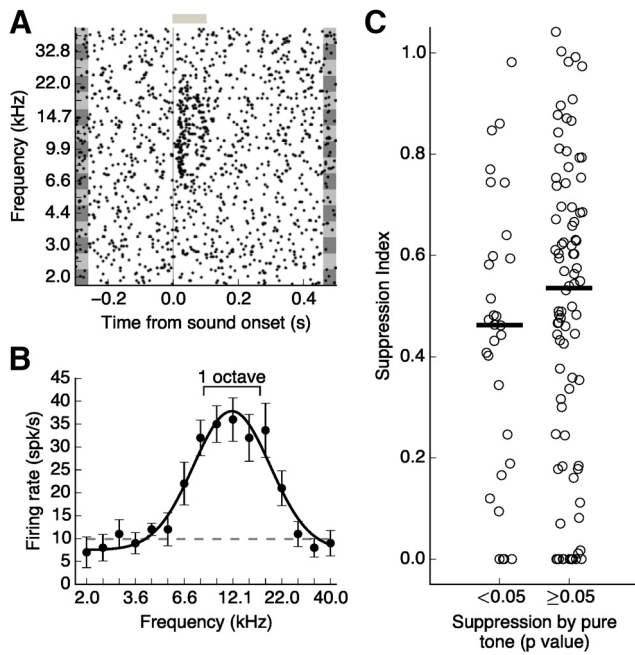
Figure 4. Surround suppression cannot be explained by suppression by PTs. **A**, Response of the putative excitatory cell depicted in Figure 2C to PTs of different frequencies. Gray bar represents sound duration (100 ms). **B**, Frequency tuning curve of cell shown in **A**. Error bars indicate SEM. Dashed line indicates the cell's spontaneous firing rate. Scale bar represents the range of frequencies in a sound stimulus with a bandwidth of one octave centered on the estimated preferred frequency. Although each individual frequency in this range elicits a positive response in this cell, the ensemble of these stimuli leads to a suppressed response (Figure 2C). **C**, Spectral surround suppression for all recorded putative excitatory cells ($N = 116$), split by whether any PT elicited a significant suppression below baseline ($p < 0.05$, Wilcoxon rank-sum test). Black bars represent medians. Suppression did not significantly differ between the two groups ($p = 0.298$, Wilcoxon rank-sum test).
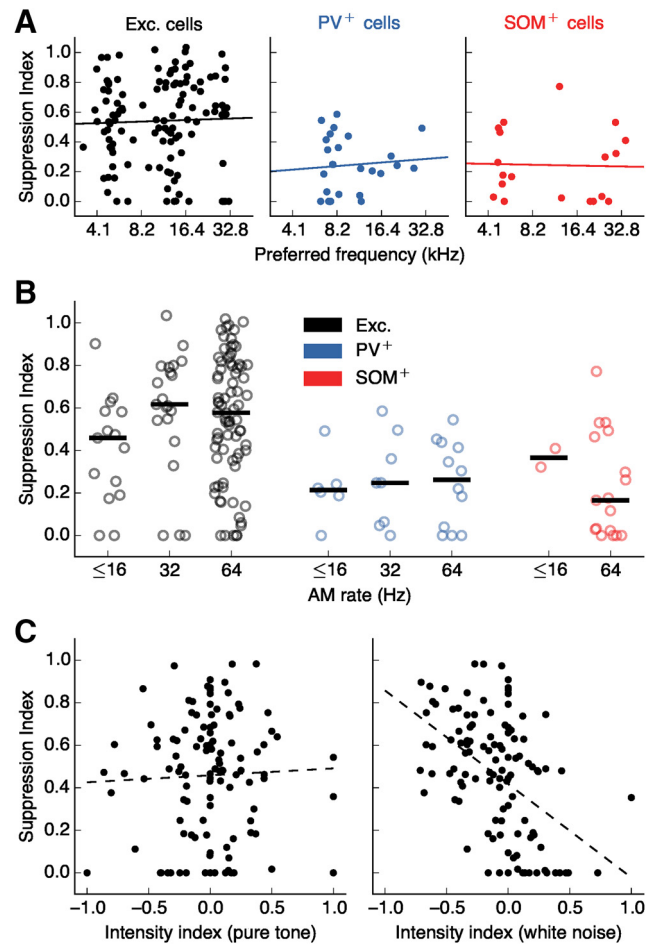

Figure 5. Surround suppression cannot be explained by other factors. **A**, SI was not correlated with preferred frequency (excitatory: $r = 0.032$, $p = 0.730$; PV$^+$: $r = 0.085$, $p = 0.681$; SOM$^+$: $r = -0.028$, $p = 0.908$). Lines indicate linear regressions. **B**, Differences in surround suppression across cell types cannot be explained by the stimulus modulation rates used (excitatory $p = 0.108$, PV$^+$ $p = 0.832$, SOM$^+$ $p = 0.350$, Kruskal–Wallis test for effect of AM rate on SI within each cell type). **C**, Left, There was no statistically significant correlation between surround suppression and intensity tuning for PTs ($r = 0.039$, $p = 0.683$). A negative intensity index means the low-intensity PT elicited a higher sustained response than the high-intensity PT. Line indicates a linear regression. Right, Intensity tuning for WN is negatively correlated with surround suppression ($r = -0.477$, $p = 9.1 \times 10^{-8}$, linear regression).

linear regression). Excitatory cells in primary auditory areas ($N = 68$) showed similar amounts of suppression to excitatory cells recorded in nonprimary auditory areas ($N = 48$; $p = 0.556$, Wilcoxon rank-sum test).

In comparison to excitatory cells, typical PV$^+$ (Fig. 2D) and SOM$^+$ (Fig. 2E) cells showed low amounts of surround suppression, as high bandwidth sounds tended to robustly drive these cells above their spontaneous firing rate. Comparing suppression indices across the populations of distinct cell types revealed that both PV$^+$ and SOM$^+$ cells exhibited less surround suppression of their sustained sound responses than excitatory cells (Fig. 2F; excitatory-PV$^+$ $p = 5.9 \times 10^{-6}$, excitatory-SOM$^+$ $p = 0.00039$, PV$^+$-SOM$^+$ $p = 0.863$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons). To verify that these differences in surround suppression across cell types could not be explained by our selection of putative excitatory cells only from SOM-ChR2 mice, we performed the same analysis after including cells that did not respond to laser in PV-ChR2 mice, which yielded a comparable magnitude of suppression in excitatory cells (SI = 0.46). Differences between excitatory and inhibitory cells were still clearly present when including this additional set of putative excitatory cells (excitatory-PV$^+$ $p = 0.0003$, excitatory-SOM$^+$ $p = 0.0056$, PV$^+$-SOM$^+$ $p = 0.863$, Wilcoxon rank-sum test). We also found that SOM$^+$ cells reached their peak response at higher bandwidths compared with putative excitatory cells (Fig. 2G; excitatory-PV$^+$ $p = 0.327$, excitatory-SOM$^+$ $p = 0.0083$, PV$^+$-SOM$^+$ $p = 0.260$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons).

Our initial characterization of responses to increasing stimulus bandwidth relied on presenting bandpass-filtered noise, centered on the cell's preferred frequency, from 0.25 octaves to WN. We did not include PTs (sinusoidal waves) in this analysis as these stimuli were generated differently. However, because much work in the field of audition has relied on the presentation of PTs, we also evaluated what happens when PTs, which have zero bandwidth, are included in the bandwidth tuning curves and the analysis of surround suppression. We found that all effects described above are present when we include responses to PTs at each cell's preferred frequency (Fig. 3). Specifically, excitatory cell responses were still more strongly suppressed at high bandwidths than the responses of PV$^+$ and SOM$^+$ cells (Fig. 3B; excitatory-PV$^+$ $p = 1.9 \times 10^{-7}$, excitatory-SOM$^+$ $p = 5.3 \times 10^{-5}$, PV$^+$-SOM$^+$ $p = 0.918$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons). Furthermore, a direct comparison between the responses to PTs and WN showed that excitatory cells have a higher preference for PT stimuli than PV$^+$ or SOM$^+$ cells, consistent with their suppression of responses to high bandwidth stimuli (Fig. 3C; excitatory-PV$^+$
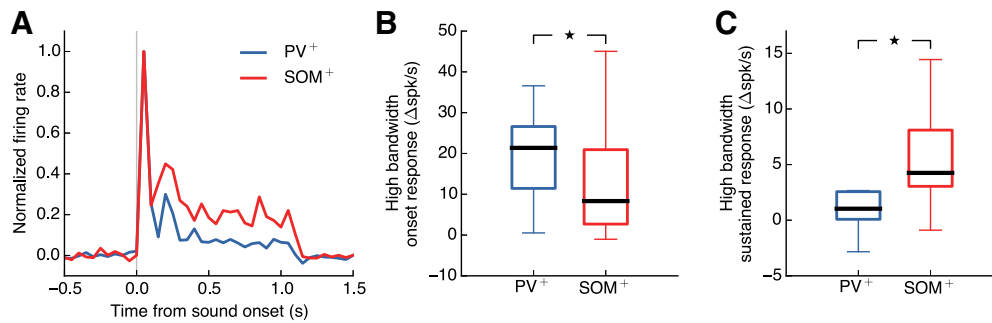
Figure 6. SOM$^+$ cells exhibit stronger sustained responses to high bandwidth stimuli than PV$^+$ cells. **A**, Median response across all PV$^+$ cells (blue) and SOM$^+$ cells (red) included in Figure 2 to high bandwidth sound stimuli ($\geq$4 octaves), baseline-subtracted and normalized so response to sound onset for both cell types is 1. Responses of SOM$^+$ cells do not drop off after onset to the same extent as PV$^+$ cell responses. **B**, Onset responses to high bandwidth sound stimuli ($\geq$ 4 octaves) are significantly higher for PV$^+$ cells than SOM$^+$ cells ($p = 0.027$, Wilcoxon rank-sum test). Black bars represent medians. Boxes extend from the lower to the upper quartile. Whiskers indicate the range. **C**, Sustained responses to high bandwidth sound stimuli ($\geq$4 octaves) are significantly higher for SOM$^+$ cells than PV$^+$ cells ($p = 0.030$, Wilcoxon rank-sum test). $*p < 0.05$.

$p = 0.0088$, excitatory-SOM$^+$ $p = 0.0026$, PV$^+$-SOM$^+$ $p = 0.113$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons). Additionally, we still observed that SOM$^+$ cells prefer significantly higher bandwidths than excitatory cells (Fig. 3D; excitatory-PV$^+$ $p = 0.920$, excitatory-SOM$^+$ $p = 0.0044$, PV$^+$-SOM$^+$ $p = 0.056$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons). Consistent with their preference for higher bandwidth sounds, we found that the majority of SOM$^+$ cells preferred WN to PTs (Fig. 3C). The majority of all cells showed some degree of increase in firing rate as bandwidths increased from PT, with no significant difference in this facilitation across cell types (Fig. 3E; excitatory-PV$^+$ $p = 0.187$, excitatory-SOM$^+$ $p = 0.498$, PV$^+$-SOM$^+$ $p = 0.089$, Wilcoxon rank-sum test, Bonferroni correction for multiple comparisons). These results indicate that the small percentage of cells (10%-19% depending on cell type) that respond most strongly to PTs and more weakly at all nonzero bandwidths do not affect the trends observed in the entire cell population.

Analysis of onset responses (0–50 ms after sound onset) of excitatory cells indicated a much weaker surround suppression during this period compared with that of the sustained responses (median excitatory onset SI: 0.25; median excitatory sustained SI: 0.49; $p = 0.0001$, Wilcoxon rank-sum test). Differences in surround suppression across some cell types were still present during this period (excitatory-PV$^+$ $p = 0.025$, excitatory-SOM$^+$ $p = 0.774$, PV$^+$-SOM$^+$ $p = 0.075$, Wilcoxon rank-sum test). Because of the much stronger spectral surround suppression observed during the sustained sound responses of excitatory cells, the remainder of our study focused on the mechanisms responsible for suppression during this period.

The strong sustained suppression brought about by high bandwidth stimuli, such as that observed for the example cell depicted in Figure 2C, can occur even when responses to brief PT stimuli outside the classical receptive field are only minimally below the spontaneous firing rate (Fig. 4A,B). Furthermore, suppression of sound responses can be present even for bandwidths encompassing frequencies for which the cell still has strong, positive responses (Fig. 4A,B). This effect was consistent across the population of putative excitatory cells, and we did not observe any differences in surround suppression between cells that did and did not exhibit significant suppression from PT stimuli (Fig. 4C; $p = 0.298$, Wilcoxon rank-sum test). This implies that the brief PT stimuli commonly used to characterize sound responses of auditory cortical neurons are not sufficient to predict the surround suppression observed here.

The observed sustained surround suppression in each cell was not correlated with the cell's preferred frequency (Fig. 5A). In addition, the observed differences in suppression across cell types cannot be explained by differences in the AM rate of the stimulus used in each case (Fig. 5B). Furthermore, while several putative excitatory cells exhibited sound intensity tuning, sometimes responding more strongly to stimuli of lower intensity, we found that the effects of intensity on evoked responses were independent from the phenomenon of surround suppression (Fig. 5C). Specifically, while neurons with high suppression indices were likely to have a preference for low-intensity WN ($r = -0.477$, $p = 9.1 \times 10^{-8}$, linear regression), there was no correlation between SI and intensity preference for PT stimuli ($r = 0.039$, $p = 0.683$, linear regression), indicating that surround suppression cannot be predicted from intensity preference in the absence of stimuli from the spectral surround.

SOM$^+$ cells exhibit stronger sustained sound-evoked responses than PV$^+$ cells

The lack of strong suppression of responses to high bandwidth stimuli observed in PV$^+$ and SOM$^+$ cells is consistent with these inhibitory cells mediating sustained suppression from the spectral surround. Analysis of the temporal dynamics of PV$^+$ and SOM$^+$ responses revealed that the responses of PV$^+$ cells were stronger at sound onset when compared with responses of SOM$^+$ cells (Fig. 6B; $p = 0.027$, Wilcoxon rank-sum test). In contrast, SOM$^+$ cells displayed stronger sustained responses than PV$^+$ cells (Fig. 6C; $p = 0.030$, Wilcoxon rank-sum test). These results imply that SOM$^+$ cells provide greater sustained inhibition in response to sounds of high bandwidth compared with PV$^+$ cells. We hypothesized, therefore, that inhibition from SOM$^+$ cells plays a dominant role in mediating spectral surround suppression in excitatory neurons.

SOM$^+$ inactivation, but not PV$^+$ inactivation, reduces spectral surround suppression in excitatory cells

To test the hypothesis that SOM$^+$ cells contribute more than PV$^+$ cells to surround suppression in excitatory cells, we recorded extracellularly from putative excitatory cells in the auditory cortex of awake mice while optogenetically inactivating either PV$^+$ cells (Fig. 7A) or SOM$^+$ cells (Fig. 7F) during sound presentation. To express the light-driven outward proton pump ArchT (Han et al., 2011) in these inhibitory cell populations, we injected a Cre-dependent adeno-associated virus (AAV5-
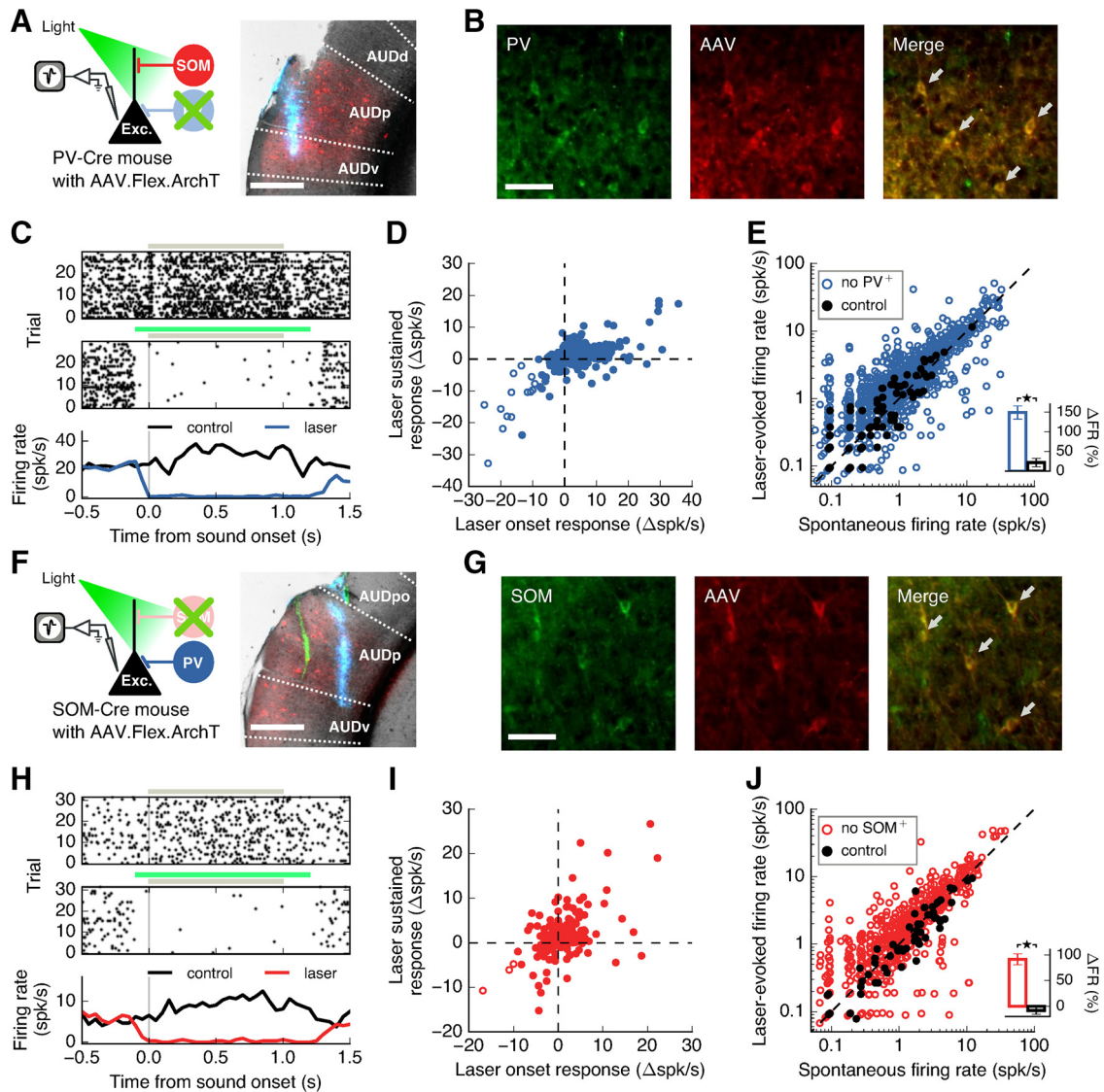
Figure 7. Inactivation of inhibitory cells in the auditory cortex. **A**, Left, Recordings from auditory cortex during photoinactivation of PV$^+$ cells. Right, Coronal brain slice (AP: −2.35 mm) showing an electrode track (blue, DiD) and the extent of virus expression at the recording location (red, tdTomato) in a PV-Cre mouse. Scale bar, 500 μm. AUDp, Primary; AUDd, dorsal; AUDv, ventral; AUDpo, posterior auditory areas. **B**, Confirmation of transfection specificity. Cells labeled with an antibody against PV (left) expressed ArchT-tdTomato (center) as indicated by the arrows (right). Scale bar, 50 μm. **C**, Response of a cell from a PV-Cre mouse to a 0.25 octave sound stimulus alone (top), or paired with laser (middle). Gray bars represent sound duration. Green bar represents laser duration. This cell is strongly inactivated by the laser. **D**, Change in firing rates of all sound-responsive cells from PV-Cre mice (N = 543 cells) at laser onset (first 50 ms) versus during sustained laser presentation (300-1100 ms after laser onset). The cells whose responses were significantly suppressed at laser onset (open dots, N = 16 cells) remained suppressed for the duration of the laser stimulus. **E**, Laser-induced firing rates (50 ms before sound onset) of single units were higher than spontaneous firing rates when laser was directed at the auditory cortex (blue points, N = 1324 cells, p = 2.3 × 10^{-54}, Wilcoxon signed-rank test), but not when laser was directed away (black points, N = 147 cells, p = 0.064, Wilcoxon signed-rank test). Inset, Directing the laser at the auditory cortex led to a stronger change in firing rate compared with the control condition (p = 3.5 × 10^{-8}, Wilcoxon rank-sum test). Error bars indicate SEM. **F**, Left, Recordings from auditory cortex during photoinactivation of SOM$^+$ cells. Right, Coronal brain slice (AP: −3.18 mm) showing electrode tracks (blue represents DiD; green represents DiO) and the extent of virus expression at the recording location (red, tdTomato) in a SOM-Cre mouse. Scale bar, 500 μm. **G**, Cells labeled with an antibody against SOM (left) expressed ArchT-tdTomato (center) as indicated by the arrows (right). Scale bar, 50 μm. **H**, Same as in **C** for an example cell from a SOM-Cre mouse. **I**, Same as in **D**, but for all sound-responsive cells from SOM-Cre mice (N = 373). Cells showing significant suppression at laser onset (open dots, N = 3) remained suppressed for the duration of the laser stimulus. **J**, Same as in **E** for SOM-Cre mice. Laser-induced firing rates were higher than spontaneous firing rates when laser was directed at the auditory cortex (red points, N = 809 cells, p = 1.0 × 10^{-30}, Wilcoxon signed-rank test), but lower in the control condition (black points, N = 78 cells, p = 0.015, Wilcoxon signed-rank test). Inset, Directing the laser at the auditory cortex led to a stronger change in firing rate compared with the control condition (p = 3.9 × 10^{-9}, Wilcoxon rank-sum test). *p < 0.05.

Flex-ArchT-tdTomato) in either PV-Cre or SOM-Cre mice. Expression of the virus spanned the majority of the primary auditory area (median range of transfection: −3.45 mm to −2.05 mm posterior from bregma; range of primary auditory area: −3.58 mm to −1.65 mm posterior from bregma). All recordings were performed in transfected areas (median range of recordings: −3.08 mm to −2.35 mm posterior from bregma; for examples, see Fig. 7A,F). Immunohistochemical analysis revealed that, in

PV-Cre mice, an anti-PV antibody labeled 84% of transfected cells, and in SOM-Cre mice, an anti-SOM antibody labeled 90% of transfected cells (Fig. 7B,G). Presentation of the laser rapidly and reliably reduced the responses of a subset of auditory cortical neurons (Fig. 7C,H). When analyzing the responses of cells showing immediate laser-induced suppression of firing rates, we found that they remained suppressed for the duration of the laser stimulus (Fig. 7D,I), implying that the laser power
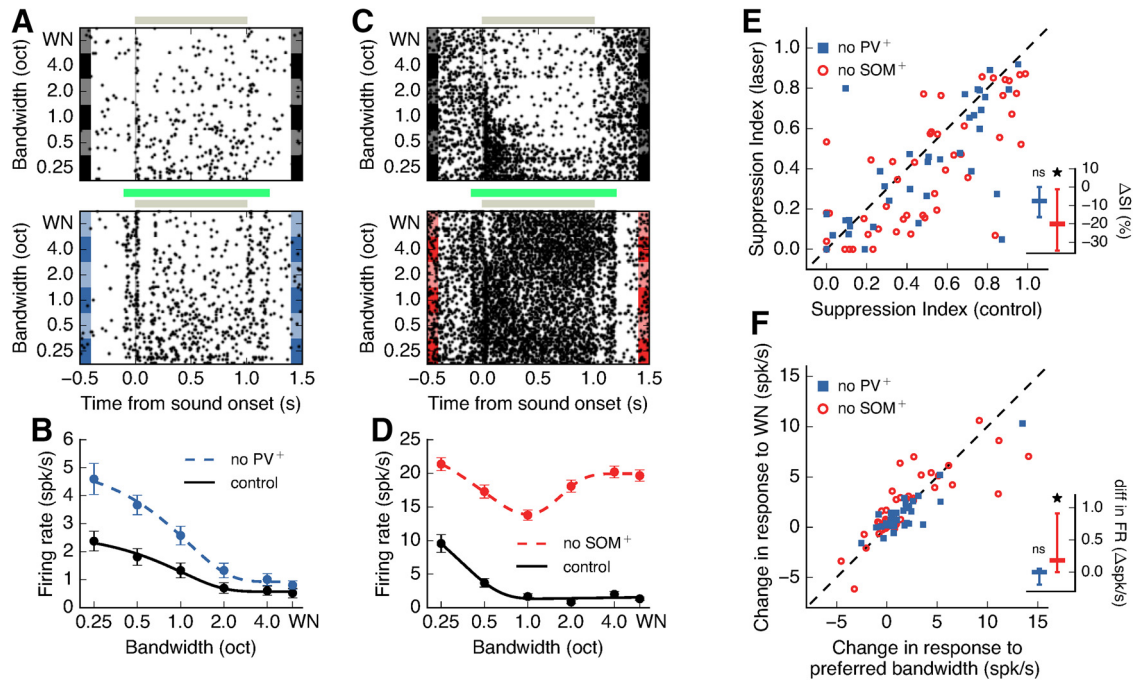
Figure 8. SOM[+] cells contribute to spectral surround suppression to a greater extent than PV[+] cells. **A**, Responses of an example putative excitatory cell to sounds of different bandwidths in the control condition (top) or during PV[+] inactivation (bottom). Gray bar represents sound duration. Green bar represents laser duration. **B**, Bandwidth tuning curve of sustained responses of example cell from **A**. Error bars indicate SEM PV[+] inactivation results in a negligible change in suppression (control SI = 0.75, PV[+]-inactivated SI = 0.79). **C**, **D**, Same as in **A**, **B** for a different putative excitatory cell during SOM[+] inactivation, which results in a large change in suppression due to increased responses to high bandwidth stimuli (control SI = 0.84, SOM[+]-inactivated SI = 0.07). **E**, Inactivation of SOM[+] cells significantly reduces surround suppression in excitatory cells ($N = 51$ cells, $p = 0.0013$, Wilcoxon signed-rank test), whereas inactivation of PV[+] cells does not ($N = 40$ cells, $p = 0.057$, Wilcoxon signed-rank test). Dashed line indicates equality. Points below this line indicate a reduction in suppression during the laser condition. Inset, Percent change in suppression evoked by laser (laser SI − control SI). *$p < 0.05$. **F**, Firing rate for high bandwidth stimuli (WN) increases more than firing rate for stimuli of the preferred bandwidth during SOM[+] inactivation ($p = 0.015$, Wilcoxon signed-rank test). The effects of PV[+] inactivation do not significantly differ by bandwidth ($p = 0.427$, Wilcoxon signed-rank test). Dashed line indicates equality. Points above this line indicate a larger increase in firing rate for WN. Four points above the dashed line from SOM[+] inactivation are outside the plot range. Inset, Difference in change in firing rate evoked by laser between the WN and preferred bandwidth conditions. A positive difference indicates a larger effect during the WN condition; ns (not significant) *$p < 0.05$.

used was sufficient to inactivate cells expressing ArchT and overcome any paradoxical network effects that may arise from inactivating inhibitory cells. Moreover, laser presentation had a significantly higher effect on firing rate when directed at the auditory cortex than when directed away (Fig. 7E,J), indicating that our results cannot be solely explained by visual stimulation from the laser.

Fig. 8A,B shows the sound responses of a typical surround-suppressed putative excitatory cell with and without inactivation of PV[+] cells. Despite the overall increase in sound-evoked firing rate, the cell maintained its previous degree of suppression (control SI = 0.75, PV[+]-inactivated SI = 0.79). In contrast, SOM[+] inactivation led to a reduction in the degree of suppression in another typical putative excitatory cell (Fig. 8C,D; control SI = 0.84, SOM[+]-inactivated SI = 0.07). In our sample of putative excitatory neurons, we found that SOM[+] inactivation led to a statistically significant reduction in sustained suppression (20% reduction on average), whereas PV[+] inactivation led to a small reduction (7.6% on average) that did not reach statistical significance (Fig. 8E; SOM[+] inactivation $p = 0.0013$, PV[+] inactivation $p = 0.057$, Wilcoxon signed-rank test). To ensure that the observed reduction in suppression was not a result of the larger sample of excitatory neurons measured during SOM[+] inactivation, we quantified the same effect after downsampling our SOM[+] inactivation data to match the sample size of the PV[+] inactivation data. Running this analysis 1000 times yielded a statistically significant effect when inactivating SOM[+] cells 99.5% of

the time, supporting the validity of the result. These results remained consistent when responses to PT stimuli were included in our estimation of surround suppression during inhibitory cell inactivation (SOM[+] inactivation $p = 0.00032$, PV[+] inactivation $p = 0.248$, Wilcoxon signed-rank test). Analysis of onset responses also showed that neither PV[+] nor SOM[+] inactivation had any significant effect on suppression of onset sound responses (PV[+] inactivation $p = 0.304$, SOM[+] inactivation $p = 0.339$, Wilcoxon signed-rank test). Furthermore, the effect of PV[+] and SOM[+] inactivation on suppression was not correlated with cortical depth (PV[+] inactivation: $r = 0.080$, $p = 0.696$; SOM[+] inactivation: $r = 0.217$, $p = 0.178$; linear regression).

We then tested whether the observed reduction in surround suppression due to SOM[+] inactivation was the result of a decrease in peak response, or an increase in the response to high bandwidth stimuli (WN), as only the latter would imply that SOM[+] cells provide information about the surround. We found that, during SOM[+] inactivation, the response to WN increased more than the response to the cell's preferred bandwidth, resulting in a reduction in surround suppression (Fig. 8F; no SOM[+] $p = 0.015$, Wilcoxon signed-rank test). In contrast, the effect of PV[+] inactivation did not differ by bandwidth ($p = 0.427$, Wilcoxon signed-rank test). These results indicate that SOM[+] cells mediate global integration of frequency information by inhibiting excitatory neural responses when spectrally distant frequencies of sounds are present.
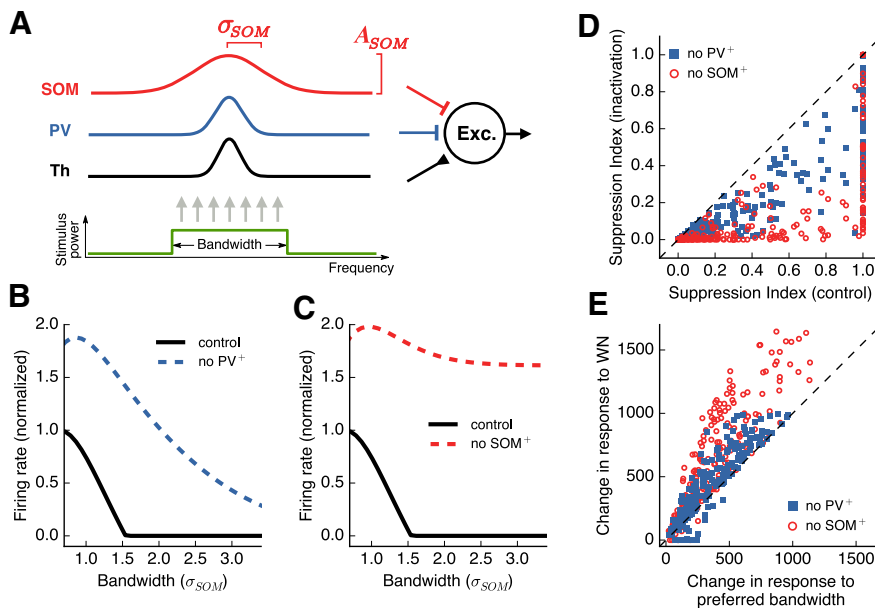
Figure 9. A feedforward model accounts for the distinct roles played by inhibitory cell types in surround suppression. **A**, A simulated output excitatory neuron receives three types of inputs weighted by Gaussian profiles. Integration across frequency channels for the SOM$^+$ input is twice as wide as the integration for PV$^+$ input. **B**, Responses of the output neuron for stimuli of different bandwidths. Inactivation of PV$^+$ input results in an increase in responses without major changes in surround suppression. **C**, Inactivation of SOM$^+$ input results in a large decrease in surround suppression. **D**, SI for each simulated output neuron, with and without inactivation of inhibitory inputs. Change in SI is consistently larger for SOM$^+$ inactivation. **E**, Change in response to preferred bandwidth and to WN stimuli, with and without inactivation of inhibitory inputs (values expressed in arbitrary model response units). The change in response to WN stimuli is consistently larger than the change for preferred bandwidth during SOM$^+$ inactivation, but not PV$^+$ inactivation.

Differences in spectral integration among cell classes accounts for their distinct roles in surround suppression

To systematically explore which features of the inputs received by excitatory cells account for the surround suppression observed in our experiments, we developed a simple, proof-of-concept model for the response tuning of excitatory cortical neurons. We modeled the response tuning of excitatory neurons given the tuning of the various inputs they receive, while ignoring several features of cortical circuitry, such as recurrent cortical and thalamocortical interactions, and spectral surround suppression in the feedforward thalamic input to the cortex. Specifically, we simulated the responses of a cortical excitatory model neuron which receives three types of inputs: (1) excitation from other cells (including thalamic inputs), (2) inhibition from PV$^+$ cells, and (3) inhibition from SOM$^+$ cells (Fig. 9A). Results from awake mice have shown that SOM$^+$ cells have broader frequency tuning than PV$^+$ or excitatory cells (Kato et al., 2017). These results suggest that SOM$^+$ cells integrate information across a wider range of frequency channels than PV$^+$ or excitatory cells, consistent with our findings when we characterized bandwidth tuning in these cell types (Fig. 2G). Given the observed differences in tuning among inhibitory cell types, we set the width of the kernel of integration for SOM$^+$ inhibition (σ_{SOM}) to be twice as large as that of PV$^+$ inhibition (σ_{PV}), and tested whether this feature was sufficient to account for the differential roles in spectral surround suppression played by these cell types.

In response to stimuli of different bandwidths, this simple model replicated the observed spectral surround suppression observed in the auditory cortex (Fig. 9B, black line). To model our inactivation experiments, we eliminated the input from each inhibitory cell class while keeping all other inputs unperturbed. The results from this simulated inactivation of inhibitory cell

classes captured the main observations of our experiments in which PV$^+$ inactivation yields an increase in activity but no major change in spectral surround suppression (Fig. 9B), while SOM$^+$ inactivation yields a major change in surround suppression (Fig. 9C).

To further test whether the effects observed in our simulations were the result of differences in spectral integration between input types and not the values chosen for other model parameters, we simulated multiple neurons, changing parameters, such as the breadth of excitation and the maximum strength of each input type, but keeping the breadth of inhibition from SOM$^+$ cells to be larger than from PV$^+$ cells. These simulations showed that the effect of inactivating SOM$^+$ cells on surround suppression is consistently larger than the effect of inactivating PV$^+$ cells, as measured by changes in SI (Fig. 9D). Moreover, the model accounted for the observation that changes in responses to WN during SOM$^+$ inactivation are much larger than changes in responses at the preferred bandwidth, an effect not observed during PV$^+$ inactivation (Fig. 9E). These simulations suggest that differences in spectral integration between PV$^+$ and SOM$^+$ cells can account for their distinct roles in surround suppression.

Discussion

In this study, we investigated how different sources of cortical inhibition contribute to spectral surround suppression, and by extension, which cortical circuits facilitate global integration over a large range of sound frequencies. By comparing the responses of PV$^+$ and SOM$^+$ cells with long broadband sound stimuli, we found that SOM$^+$ cells maintain high rates of sustained activity, whereas PV$^+$ cells instead have strong responses locked to the sound onset. Consistent with previous studies (Blackwell and Geffen, 2017; Liang et al., 2019), these results suggest that PV$^+$ cells provide fast, transient inhibition, whereas SOM$^+$ cells provide sustained inhibition. Combined with the small surround suppression observed in these inhibitory cells, our results imply that SOM$^+$ cells provide constant information to excitatory cells about the presence of frequencies in their surround. While PV$^+$ cells display the same lack of suppression seen in SOM$^+$ cells, only inactivation of SOM$^+$ cells influenced sustained surround suppression, consistent with the temporal dynamics of their evoked responses. Inactivation of PV$^+$ cells led to a small and not statistically significant reduction in suppression. This suggests that, although PV$^+$ cells do participate in the general suppression of sound responses, their contribution to the phenomenon of surround suppression at large bandwidths is comparatively small, possibly as a result of their low sustained firing rates. While both PV$^+$ and SOM$^+$ cells had relatively high firing rates at sound onset, neither was found to contribute significantly to the suppression of onset sound

responses in excitatory cells, suggesting that this effect may not be mediated by cortical inhibition, but rather inherited from upstream auditory areas.

In our characterization of surround suppression across cell types, the majority of PV[+] and SOM[+] cells showed small amounts of suppression in their responses to high bandwidth stimuli. Despite these responses in inhibitory neurons, many excitatory cells exhibited a monotonic decrease in firing rate as the bandwidth of the sound stimulus increased past their pre-ferred bandwidth. While our inactivation experiments showed that spectral surround suppression is primarily mediated by inhi-bition from SOM[+] cells, they do not fully explain the relation-ship between the activity we observe in excitatory and SOM[+] cells. The lack of reduced suppression in excitatory responses at high bandwidths, despite the reduction in inhibitory cell activity, could be explained by complex effects resulting from the recur-rence of the cortical network, or by excitatory neurons having reached their lowest level of firing at which small changes in in-hibition have little influence. Further study is needed to evaluate these hypotheses.

Our results indicate that SOM[+] cells play a dominant role in mediating spectral surround suppression of sustained sound responses by providing excitatory cells with informa-tion about their surround. This shows that SOM[+] cells play a role in modulating auditory cortical activity based on the spectral content of sounds, similar to previous findings show-ing that SOM[+] cells modulate sound responses depending on temporal context (Natan et al., 2017). Consistent with our findings, inhibition from SOM[+] cells has been shown to mediate long-range lateral inhibition evoked by PT stimuli by suppressing excitatory and inhibitory synaptic inputs onto layer 2/3 cells (Kato et al., 2017). However, we found that cells may exhibit surround suppression, even when they are not suppressed below baseline by isolated PTs outside their receptive fields, providing evidence of the nonlinearity in the interactions between a cell's response to stimuli within and outside its receptive field.

To explore the characteristics of SOM[+] cells that explain their dominant role in producing surround-suppressed responses to sound, we created a simple feedforward model of excitatory cell responses based on the frequency tuning of the excitatory and in-hibitory inputs onto these cells. There exists conflicting informa-tion about the tuning properties of excitatory, PV[+], and SOM[+] cells from the auditory cortex, with some studies finding wider tuning in PV[+] cells than excitatory and SOM[+] cells (L. Y. Li et al., 2015; Liang et al., 2019), others finding similar tuning in PV[+] cells and excitatory cells (Moore and Wehr, 2013), and others finding broader tuning in SOM[+] cells compared with other cell types (Kato et al., 2017). Although the latter study used cal-cium imaging as opposed to electrophsyiology, their use of awake animals and long-duration sound stimuli made their experimental conditions most similar to ours. We therefore chose to base our model parameters on their findings. While our choice of kernel widths for PV[+] and SOM[+] cells may seem to differ from experimentally derived connection pat-terns, which show that the spatial extents of the connections between excitatory and inhibitory cells are roughly equal for all cell types (Levy and Reyes, 2012), our kernels represent a combination of the tuning width of presynaptic responses and connection ranges. As such, the broad frequency tuning observed in SOM[+] cells is sufficient to justify our choice of kernels. Furthermore, studies show that SOM[+] cells influence excitatory cell activity at a longer range than PV[+] cells (L. Y.

Li et al., 2014; Kato et al., 2017), providing another possible explanation for the inhibition by spectrally distant stimuli provided by SOM[+] cells.

Despite the simplicity of our model, the broad tuning of inhibition from SOM[+] cells was sufficient to explain the effects on surround suppression we found when inactivating SOM[+] cells, consistent with network simulations showing that lateral inhibition can arise from broadly tuned inhibition (Levy and Reyes, 2011). However, this simple model does not fully account for other phenomena present in our data, such as the surround suppression (albeit small) observed in inhibi-tory cell responses, or the fact that a subset of cells become more surround-suppressed with the removal of inhibitory inputs, leaving open an opportunity for future improvements. The inclusion of recurrent interactions across cells, surround suppression of feedforward thalamic inputs, and saturation of inhibitory inputs in the model could potentially address these shortcomings.

The phenomenon of surround suppression is also present in sensory systems other than the auditory system. In the visual cor-tex, responses to visual stimuli of most excitatory cells in layer 2/ 3 decrease as the stimulus grows in size. This suppression was found to be mediated by inhibition from SOM[+] cells (Adesnik et al., 2012), consistent with our results from the auditory cortex. However, our study found that the strength of sustained responses and degree of suppression seen in PV[+] cells in the au-ditory cortex do not fully match those found in the visual cortex. These observations suggest that, while the role of each cell type may not be fully replicated throughout the cortex, inhibition from SOM[+] cells is a common mechanism for producing responses that are suppressed by the sensory surround.

While our results show that the broad category of SOM[+] cells plays a role in spectral surround suppression, SOM[+] cells can be further classified into distinct subtypes by their anatomy and gene expression patterns (Ma et al., 2006). Recent work in the somatosensory system demonstrated that two different subsets of SOM[+] cells, Martinotti and non-Martinotti, specifically and reciprocally interconnect with excitatory cells in different cortical layers (Naka et al., 2019), and as such may contribute differently to surround suppression. While the specific contributions of each cell type have yet to be determined, our results demonstrate that these types of SOM[+] cells, either individually or in tandem, are key in shaping the responses of surround-suppressed audi-tory cortical excitatory cells.

Our results reveal that a large proportion of excitatory cells in the auditory cortex exhibit suppression of activity by their spec-tral surround. As a result of this phenomenon, neural responses to broadband stimuli, such as many naturally occurring types of acoustic noise, are reduced. We therefore posit that spectral sur-round suppression of activity can be used by the auditory system to filter out background noise while transmitting narrowband signals. Our study identified a role of SOM[+] cells in the imple-mentation of this integration of acoustic features for noise reduc-tion, and provides evidence of common mechanisms across sensory systems.

References

Abeles M, Goldstein MH (1972) Responses of single units in the primary au-ditory cortex of the cat to tones and to tone pairs. Brain Res 42:337–352.

Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M (2012) A neural circuit for spatial summation in visual cortex. Nature 490:226–231.

Ayaz A, Saleem A, Schölvinck M, Carandini M (2013) Locomotion controls spatial integration in mouse visual cortex. Curr Biol 23:890–894.

Blackwell JM, Geffen MN (2017) Progress and challenges for understanding the function of cortical microcircuits in auditory processing. Nat Commun 8:2165.

Han X, Chow BY, Zhou H, Klapoetke NC, Chuong A, Rajimehr R, Yang A, Baratta MV, Winkle J, Desimone R, Boyden ES (2011) A high-light sensitivity optical neural silencer: development and application to optogenetic control of non-human primate cortex. Front Syst Neurosci 5:18.

Kadir SN, Goodman DF, Harris KD (2014) High-dimensional cluster analysis with the masked EM algorithm. Neural Comput 26:2379–2394.

Kato HK, Asinof SK, Isaacson JS (2017) Network-level control of frequency tuning in auditory cortex. Neuron 95:412–423.e4.

Kuchibhotla KV, Gill JV, Lindsay GW, Papadoyannis ES, Field RE, Sten TA, Miller KD, Froemke RC (2017) Parallel processing by cortical inhibition enables context-dependent behavior. Nat Neurosci 20:62–71.

Kvitsiani D, Ranade S, Hangya B, Taniguchi H, Huang JZ, Kepecs A (2013) Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature 498:363–366.

Levy RB, Reyes AD (2011) Coexistence of lateral and co-tuned inhibitory configurations in cortical networks. PLoS Comput Biol 7:e1002161.

Levy RB, Reyes AD (2012) Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. J Neurosci 32:5609–5619.

Li H, Liang F, Zhong W, Yan L, Mesik L, Xiao Z, Tao HW, Zhang LI (2019) Synaptic mechanisms for bandwidth tuning in awake mouse primary auditory cortex. Cereb Cortex 29:2998–3009.

Li LY, Ji XY, Liang F, Li YT, Xiao Z, Tao HW, Zhang LI (2014) A feedforward inhibitory circuit mediates lateral refinement of sensory representation in upper layer 2/3 of mouse primary auditory cortex. J Neurosci 34:13670–13683.

Li LY, Xiong XR, Ibrahim LA, Yuan W, Tao HW, Zhang LI (2015) Differential receptive field properties of parvalbumin and somatostatin inhibitory neurons in mouse auditory cortex. Cereb Cortex 25:1782–1791.

Liang F, Li H, Chou X, Zhou M, Zhang NK, Xiao Z, Zhang KK, Tao HW, Zhang LI (2019) Sparse representation in awake auditory cortex: cell-type dependence, synaptic mechanisms, developmental emergence, and modulation. Cereb Cortex 29:3796–3812.

Ma Y, Hu H, Berrebi AS, Mathers PH, Agmon A (2006) Distinct subtypes of somatostatin-containing neocortical interneurons revealed in transgenic mice. J Neurosci 26:5069–5082.

Metherate R, Kaur S, Kawai H, Lazar R, Liang K, Rose HJ (2005) Spectral integration in auditory cortex: mechanisms and modulation. Hear Res 206:146–158.

Moore AK, Wehr M (2013) Parvalbumin-expressing inhibitory interneurons in auditory cortex are well-tuned for frequency. J Neurosci 33:13713–13723.

Naka A, Veit J, Shababo B, Chance RK, Risso D, Stafford D, Snyder B, Egladyous A, Chu D, Sridharan S, Mossing DP, Paninski L, Ngai J, Adesnik H (2019) Complementary networks of cortical somatostatin interneurons enforce layer specific control. eLife 8:e43696.

Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM, Geffen MN (2015) Complementary control of sensory adaptation by two types of cortical interneurons. eLife 4:e09868.

Natan RG, Rao W, Geffen MN (2017) Cortical interneurons differentially shape frequency tuning following adaptation. Cell Rep 21:878–890.

Nelken I, Prut Y, Vaadia E, Abeles M (1994) Population responses to multifrequency sounds in the cat auditory cortex: one- and two-parameter families of sounds. Hear Res 72:206–222.

O'Connor KN, Yin P, Petkov CI, Sutter ML (2010) Complex spectral interactions encoded by auditory cortical neurons: relationship between bandwidth and pattern. Front Syst Neurosci 4:145.

Phillips EA, Schreiner CE, Hasenstaub AR (2017) Cortical interneurons differentially regulate the effects of acoustic context. Cell Rep 20:771–778.

Rauschecker JP, Tian B (2004) Processing of band-passed noise in the lateral auditory belt cortex of the Rhesus monkey. J Neurophysiol 91:2578–2589.

Rudy B, Fishell G, Lee S, Hjerling-Leffler J (2011) Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol 71:45–61.