HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2017 March 02.

Published in final edited form as:

Neuron. 2016 March 2; 89(5): 1031–1045. doi:10.1016/j.neuron.2016.01.027.

Cross-modality Sharpening of Visual Cortical Processing through Layer 1-Mediated Inhibition and Disinhibition

Leena A. Ibrahim[1,4], **Lukas Mesik**[1,4], **Xu-ying Ji**[1], **Qi Fang**[1,4], **Hai-fu Li**[1], **Ya-tang Li**[1], **Brian Zingg**[1,4], **Li I. Zhang**[1,3], and **Huizhong Whit Tao**[1,2]

[1]Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, CA 90033, USA

[2]Department of Cell and Neurobiology, University of Southern California, Los Angeles, CA 90033, USA

[3]Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

[4]Neuroscience Graduate Program, University of Southern California, Los Angeles, CA 90033, USA.

Summary

Cross-modality interaction in sensory perception is advantageous for animals' survival. How cortical sensory processing is cross-modally modulated and what are the underlying neural circuits remain poorly understood. In mouse primary visual cortex (V1), we discovered that orientation selectivity of layer (L)2/3 but not L4 excitatory neurons was sharpened in the presence of sound or optogenetic activation of projections from primary auditory cortex (A1) to V1. The effect was manifested by decreased average visual responses yet increased responses at the preferred orientation. It was more pronounced at lower visual contrast, and was diminished by suppressing L1 activity. L1 neurons were strongly innervated by A1-V1 axons and excited by sound, while visual responses of L2/3 vasoactive intestinal peptide (VIP) neurons were suppressed by sound, both preferentially at the cell's preferred orientation. These results suggest that the cross-modality modulation is achieved primarily through L1 neuron and L2/3 VIP-cell mediated inhibitory and disinhibitory circuits.

Keywords

Primary visual cortex; *in vivo* patch-clamp recording; Ca^{2+} imaging; auditory cortex; cross-modal interaction; interneuron

Correspondence should be addressed to H.W. Tao: ; Email: htao@usc.edu, or L.I. Zhang: ; Email: liizhang@usc.edu..

Author contributions

H.W.T and L.I.Z. conceived and supervised the study. L.A.I. performed loose-patch recording experiments. L.A.I. and L.M. performed Ca^{2+} imaging experiments and analysis. X.J. performed slice recording experiments. Q.F. and H. L. performed awake recording experiments. B.Z. and L.A.I. did anatomical analysis. Y.T.L. was involved in *in vivo* recordings at the early phase of the project. H.W.T, L.I.Z. and L.A.I. wrote the manuscript.

Introduction

Animals constantly receive a multitude of sensory inputs, such as visual, auditory and somatosensory, from the outside world, which are eventually processed in the cortex. The precise processing of sensory input from each of these modalities is essential for an animal's perception, behavior and survival. Previously, it has been thought that information of different sensory modalities is first processed in anatomically isolated primary sensory areas, and that merging of different sensory information then takes place in higher association areas (Felleman and Van Essen, 1991; Jones and Powell, 1970). This view mainly stems from lesion studies showing that deficits in a specific primary sensory cortex only affect the perception of the corresponding modality (Dewson et al., 1970; Winans, 1967). In addition, earlier anatomical studies have not been able to clearly reveal cross-modality projections (KUYPERS et al., 1965).

Such view has been challenged more recently, as a number of functional magnetic resonance imaging (fMRI) studies have shown that cross-modality interactions may occur at the level of primary sensory cortices (Clavagnier et al., 2004; Kayser et al., 2005; Kok et al., 2012). For example, it has been found that the spatial pattern of activation of a primary sensory cortical area (e.g. V1) is different when stimuli from a different modality (auditory or tactile) are presented (Liang et al., 2013). This finding suggests that there may be discrete locations within V1 that respond to specific cross-modal inputs. Furthermore, an increasing number of animal studies with viral tracing in both anterograde and retrograde manners have begun to reveal patterns of potential cross-modality interactions (Campi et al., 2010; Cappe and Barone, 2005; Falchier et al., 2002; Frostig et al., 2008; Sieben et al., 2013; Stehberg et al., 2014; Zingg et al., 2014). In humans, psychophysics studies have shown that combining sensory information of different modalities is important for speeding reaction times (Gielen et al., 1983; HERSHENSON, 1962) and for detecting indistinct stimuli (Driver and Spence, 1998; Frens et al., 1995; Jaekl and Harris, 2009; McDonald et al., 2000; Vroomen and de Gelder, 2000). A recent behavioral study in rats has also shown that performance success rate is higher and reaction time is faster when visual targets are accompanied by a simultaneous sound (Gleiss and Kayser, 2012).

While evidence is now emerging to increasingly support occurrence of multisensory integration in primary sensory cortices (Iurilli et al., 2012), the role of cross-modality inputs in modulating sensory processing properties of cortical neurons and the underlying neural circuits remain not well understood. In this study, we found that in the presence of sound or optogenetic activation of axonal projections from A1 to V1, the orientation tuning of L2/3 pyramidal neurons was sharpened. This resulted from an increase of firing rate at the preferred orientation in accompany with a general reduction of average visual responses. Recordings in brain slices revealed that L1 inhibitory neurons received the maximum amount of direct input from A1 as compared with other cell types across all the cortical layers, including parvalbumin (PV), somatostatin (SOM), and vasoactive intestine peptide (VIP) inhibitory neurons, as well as pyramidal cells. Comparisons of visual responses of different inhibitory neurons in superficial layers of V1 in the absence and presence of sound

further suggested local inhibitory and disinhibitory circuits that underlie the sound-induced modulation in visual cortex.

Results

Effects of sound stimulation on excitatory neuron responses in L2/3 of V1

Using *in vivo* cell-attached loose-patch recording, we examined spike responses of L2/3 neurons in V1 to drifting sinusoidal gratings at 12 different directions (6 orientations) and at 25% contrast in anaesthetized mice. Our recording method was highly biased towards sampling excitatory neurons (Wu et al., 2008; Liu et al., 2010). Visual stimuli in the absence and presence of white noise pulses (50 ms pulse duration, at 10 Hz, for 1.5 sec, 70 dB sound pressure level (SPL)) were interleaved, and stimuli of different directions were presented in a random fashion. As shown by an example cell in Figure 1A, sound apparently increased the evoked spike number at the cell's preferred orientation, while decreased the spike number at the orientation orthogonal to the preferred. We used a global orientation selectivity index (gOSI), which is equivalent to 1 minus circular variance (Mariño et al., 2005), to quantify the degree of orientation selectivity (OS) (see Experimental Procedures). The cell exhibited an increased gOSI in the presence of sound (Figure 1A, right panel), indicating that its OS had been sharpened.

In a total of 26 similarly recorded orientation tuned pyramidal neurons, the increase of evoked firing rate at the preferred orientation in the presence of sound was statistically significant, and firing rates at some orientations away from the preferred were significantly decreased (Figure 1B). On average, there was a 22% increase in the evoked firing rate at the preferred orientation (Figure 1C), and 32% decrease at the orthogonal orientation (Figure 1D), and 9% decrease in the overall firing rate averaged across orientations ($p < 0.01$ paired t-test). These changes led to a significant increase in gOSI (Figure 1E). The effect was smaller when orientation tuning was measured with gratings at 95% contrast (Figure S1A). Sound stimulation alone resulted in decreased spontaneous firing rates of these neurons (Figure S1B), indicating a modulatory role of sound. In contrast to L2/3, sound had no significant effect on the response level or OS of L4 pyramidal neurons (Figure S1C-S1E).

We also observed similar sound effects in the quiet wakeful state, with loose-patch recordings in un-anaesthetized animals (Figure S2). On average, there was a 24% increase of firing rate at the preferred orientation (Figure S2C), and 48% decrease at the orthogonal orientation (Figure S2D). These together led to a significant increase in gOSI ($p < 0.01$, paired t-test) (Figure S2E).

To further confirm the phenomenon with a larger number of cells, we performed two-photon Ca^{2+} imaging in L2/3 of anaesthetized mice by injecting adeno-associated virus (AAV) encoding GCaMP6 in V1. Although the non-floxed GCaMP6 in principle had labeled both excitatory and inhibitory neurons, the great majority of fluorescent cells imaged ought to be pyramidal cells due to their large abundance in the cortex. We made sequential line scanning across selected individual cell bodies so that Ca^{2+} responses of 10-20 neurons could be imaged in one experiment with a reasonably good temporal resolution (Figure 2A, left panel). The time-dependent fractional change of fluorescence intensity ($\Delta F/F0$) was

determined for each stimulus direction over a period of 10 sec. As shown by an example L2/3 neuron, an increase in ΔF/F0 was observed at the preferred orientation and a simultaneous decrease was seen at some non-optimal orientations (Figure 2A, 2B). Overall, there was a slight decrease in Ca^{2+} response averaged across orientations when sound was applied (Figure 2C).

Tuning curves generated from peak ΔF/F0s were plotted for a total of 75 L2/3 neurons (Figure 2D). We summarized data in two ways. In the first, we averaged normalized ΔF/F0s across cells for each test orientation (relative to the preferred orientation in the sound off condition), and fit the average results with Gaussian functions (Figure 2E, left panel). In the second, for each individual cell, we first fit the tuning in each condition with a Gaussian curve, and then averaged Gaussian curves of all the cells (Figure 2E, right panel). By either way of analysis, we observed an increase of Ca^{2+} response at the preferred orientation, and a decrease of the response at orientations away from the preferred. Overall, the preferred orientation of a cell included in the analysis was not significantly different between sound-off and sound-on conditions ($p > 0.05$, paired t-test, Figure 2F). There was a 10% increase in Ca^{2+} response at the preferred orientation (Figure 2G), and 15% decrease at the orthogonal orientation (Figure 2H), and 5% decrease in the average response ($p < 0.001$, paired t-test). These changes resulted in a significant increase in gOSI (Figure 2I). Similar conclusions could be made when we quantified orientation tuning with a more tradition index, OSI, which is defined by $(R_{pref}-R_{orth})/(R_{pref}+R_{orth})$ (Figure S3A-S3C). Thus, the imaging data were consistent with the loose-patch recording results, confirming the effect of sharpening OS by sound.

V1 receives direct inputs from A1

Auditory information must enter visual cortex for the sound modulation effect. A previous study has implied that A1 might make anatomical connections with V1 (Iurilli et al., 2012). Here, we traced the direct projections from A1 to V1 by injecting a retrograde tracer, CTb-Alexa 488, into V1 (Figure 3A, left panel). The retrogradely labeled cells in A1 were observed in different layers, but most densely in L5 (Figure 3A, middle and right panels). Since L5 contained the largest portion of V1 projecting neurons, we injected AAV encoding floxed humanized channelrhodopsin 2 (hChR2) fused with EYFP into A1 of Rbp4-Cre mice, a L5-specific Cre line (Figure 3B, left panel). In V1, fluorescence-labeled A1 axons terminated mainly in superficial layers of V1, with the highest density observed in L1 (Figure 3B, middle panel). The strength of this A1-V1 projection in L1 was comparable to the A1 projection to the dorsal cortex of inferior colliculus (DCIC) (Figure S4A), a well-known corticofugal target of A1 (Andersen et al., 1980; Bajo and Moore, 2005; Bajo et al., 2007; Markovitz et al., 2013; Xiong et al., 2015).

Optogenetic stimulation of A1-V1 axons sharpens OS in L2/3

After 2-3 weeks of viral expression of hChR2 in A1, we optogenetically activated A1-V1 axons. Blue LED light (470 nm) illumination was delivered through an optic fiber placed on the surface of V1 or A1 (see Experimental Procedures). Trains of light pulses (10-ms duration) were applied at 20 Hz for 1.5 sec to cover the entire duration of visual stimulation. LED stimulation alone elicited local field potential changes in superficial layers of V1

(Figure 3B, right panel), indicating the effectiveness of activation. We then examined spike responses of individual L2/3 cells to moving gratings without and with LED stimulation at two different visual contrasts (95% and 25%). Similar sharpening of OS was observed, as demonstrated by an increase and decrease in evoked firing rate at the preferred and orthogonal orientation, respectively (Figure 3C-3J). The effect was also more pronounced at the lower visual contrast (Figure 3D-3F, 3H-3J, also see Figure 4SC). The increase in firing rate at the preferred orientation was 26% at 25% contrast, and 12% at 95% contrast. The decrease in firing rate at the orthogonal orientation was 40% at 25% contrast, and 24% at 95% contrast. The increase in gOSI was 25% at 25% contrast, and 15% at 95% contrast. With the visual stimulation alone, orientation tuning was sharper at the higher contrast (gOSI = 0.37 ± 0.12 and 0.46 ± 0.13 at 25% and 95% contrast respectively, $p < 0.01$, paired t-test, $n = 10$), consistent with our previous report (Li et al., 2012a). In L4, however, stimulating A1-V1 axons did not have any significant effect on orientation tuning or evoked firing rates (Figure 3K-3R, and Figure S3D-S3E). These data suggest that particularly under low stimulus strength conditions (as with low contrast gratings), sound has great benefit to vision in that it renders output neurons in V1 to be better tuned for orientation. The absence of sound effects in L4 suggests a top-down rather than thalamocortical modulation.

Innervation pattern of the A1-V1 projection

To understand the circuit mechanism underlying the sound effects and in particular how inhibitory neurons were involved, we crossed inhibitory neuron-specific Cre lines (PV, SOM, VIP and GAD2-Cre) (Taniguchi et al., 2011) with a Cre-dependent tdTomato reporter line, Ai14, to label individual inhibitory subtypes or all inhibitory neurons. We then injected AAV encoding non-floxed hChR2 in A1, and performed whole-cell voltage clamp recordings from desired inhibitory neurons as well as pyramidal cells (morphologically identified, or identified as non-fluorescence-labeled cells in GAD2-Cre::tdTomato mice) in V1 slices (Figure 4A). To record monosynaptic excitatory currents only, TTX (1 μM) and 4-AP (1 mM) were present in the bath solution (Cruikshank et al., 2010; Ji et al., 2015; Petreanu et al., 2009). LED illumination was applied to the entire visual cortical area to activate A1 axons in this region. As shown by response traces of example cells (Figure 4B) and summary plots of relative peak response amplitude (Figure 4C, 4D, see Experimental Procedures), L1 inhibitory neurons, except those positive for VIP which only comprise ~10% of the L1 population (Rudy et al., 2011), received strong excitatory input from A1, much stronger as compared with the other cell types examined, including PV, SOM and VIP inhibitory neurons as well as pyramidal cells. In L2/3, the strength of A1 input to PV neurons was comparable to that to pyramidal cells ($p = 0.97$, two sample t-test). Both were higher than the inputs to SOM and VIP neurons ($p < 0.01$, t-test) (Figure 4D). Beyond 400-μm depth, none of the cell types received direct A1 input (Figure 4C), consistent with the *in vivo* recording data that L4 responses were not modulated by sound. Therefore, the strongest excitatory input from A1 is received by L1 inhibitory neurons (those negative for VIP), followed by excitatory neurons and PV cells in L2/3. Since all L1 neurons are inhibitory (Beaulieu et al., 1992; Chu et al., 2003; Li and Schwark, 1994; Winer and Larue, 1989; Wozny and Williams, 2011; Zhou and Hablitz, 1996; Zhou et al., 2014), these results suggest that L1 cells are in a suitable position to provide general inhibition to L2/3 pyramidal neurons, decreasing their overall visual responses.

Activation of L1 neurons by sound

To examine whether visual responses of L1 neurons were affected by sound stimulation, we performed Ca^{2+} imaging in GAD2-Cre::tdTomato mice injected with AAV encoding floxed GCaMP6s in V1 (Figure S5A). We found that most of imaged L1 neurons were robustly activated in the presence of sound as opposed to visual stimulation alone. As shown by the plots of ΔF/F0 for an example L1 neuron (Figure 5A), there was a marked increase in Ca^{2+} response level at almost every stimulus orientation (Figure 5B), resulting in a pronounced enhancement of the overall response averaged across orientations (Figure 5C). Tuning curves of normalized peak response level are plotted for 40 L1 cells (Figure 5D). In the presence of sound, we observed an increase of visually evoked Ca^{2+} response at nearly all orientations, either by averaging normalized response levels across cells (Figure 5E, left panel) or by averaging individually fitted Gaussian curves (Figure 5E, right panel). On average, there was a 20% increase in Ca^{2+} response at the preferred orientation (Figure 5F), 10% increase at the orthogonal orientation (Figure 5G), and 13% increase in the overall response level (Figure 5H). The latter increase was relatively larger at the lower visual contrast (Figure S6A), and enhanced with increasing sound intensities (Figure S6B). L1 neurons overall were tuned, although the tuning was relatively broad (Figure 5E). Since the increase in Ca^{2+} response was relatively larger at the preferred than orthogonal orientation (Figure 5F, 5G), the OS of L1 neurons was enhanced in the presence of sound (Figure 5I).

Additionally, sound stimulation alone could excite L1 neurons in V1 (Figure 6A-6B). In one experiment, we imaged Ca^{2+} responses of a group of L1 neurons and L2/3 pyramidal cells in the same animal and with the same temporal resolution, and found that visual responses in L2/3 appeared delayed relative to the sound-evoked responses in L1 (Figure 6C). To be able to more precisely compare the onset timings of these two types of responses, we performed loose-patch recordings from L1 neurons, which were easily identified by their spike responses to sound (Figure 6D, 6E). Indeed, relative to the visual responses of L2/3 pyramidal cells (Figure 6F), the onset of sound-evoked spike responses of L1 neurons was much earlier (Figure 6G). Given that sound and visual stimulation were applied simultaneously in our experiments of cross-modality interaction, this result suggests that L1 neurons may be able to modulate L2/3 responses as early as the latter start.

Optogenetic suppression of L1 neuron activity

To understand whether the sound-induced modulation of OS was mediated by V1 L1 neurons, we sought to selectively suppress their activity during sound presentation. Since there was no L1-specific Cre line available yet, we employed an iontophoresis method (see Experimental Procedures) to limit the spread of viral particles within L1 when injecting AAV encoding floxed ArchT-GFP in GAD2-Cre mice (Figure 7A). Green LED light (530 nm) illumination was applied onto the surface of V1 to inhibit ArchT-expressing L1 neurons. Using loose-patch recordings, we examined spike responses of L2/3 pyramidal cells to interleaved trials of visual stimulation only (moving gratings at 25% contrast), visual plus sound, and visual plus sound plus LED stimulation. As shown by the tuning curves of two example cells (Figure 7B) and summary of all the cells (Figure 7C), OS was enhanced in the presence of sound (compare red and blue curves), similar as observed earlier. When L1 neurons were suppressed, the sound-induced increase of response level at the preferred

orientation was reduced by approximately half (Figure 7D), the decrease of response level at the orthogonal orientation was nearly abolished (Figure 7E), and the reduction of the overall response level was prevented (Figure 7F). As a result, OS of pyramidal cells in the L1-suppression condition became significantly weaker as compared with the visual-plus-sound condition (Figure 7G). Therefore, suppressing L1 neurons could block, at least partially, the sound-induced sharpening of OS of L2/3 pyramidal cells, suggesting that L1 neurons are involved in mediating this cross-modality effect.

Effects of sound stimulation on L2/3 PV, SOM and VIP neurons

We further imaged Ca^{2+} responses of PV, SOM or VIP neurons in L2/3 of V1, in different inhibitory neuron-specific Cre mice injected with AAV-floxed-GCaMP6s (Figure S5B-S5D). Although all of these inhibitory subtypes received some amount of direct excitatory input from A1 (Figure 4D), the effects of sound on their OS and response levels were variable. For PV neurons, we did not observe significant changes in the visually evoked Ca^{2+} response at the preferred orientation or in the average response level when sound was applied (Figure 8A-8B). Similar results were obtained when we identified PV neurons based on their fast-spiking properties (Atallah et al., 2012; Li et al., 2015a; Ma et al., 2010) in one set of loose-patch recording experiments using pipette parameters that increased chances of encountering inhibitory cells (see Experimental Procedures) (Figure S7). For SOM neurons, we did not detect visually evoked Ca^{2+} responses in the majority (~90%) of cells under our current anesthetized condition, consistent with the notion that evoked firing rates of SOM neurons are in general low in anesthesia (Adesnik et al., 2012). In the small subset of SOM neurons exhibiting significant responses, we did not find a significant change in their average response level or response level at the preferred orientation (Figure 8C-8D). Therefore, neither PV nor SOM neurons appeared to contribute significantly to the sound-induced effect. For VIP neurons however, we observed a significant decrease of Ca^{2+} response at the preferred orientation (Figure 8E), and a smaller decrease in the average response level (Figure 8F). Responses of an example VIP neuron are shown in Figure 8G-8H, and summary tuning curves for all VIP neurons are shown in Figure 8I. Notably, the decrease of response level was more pronounced at the preferred than other orientations (Figure 8I), indicating that VIP neurons are suppressed by sound preferentially at their preferred orientation, in contrast to what had been observed in pyramidal cells.

Discussion

In this study, we examined how input of one sensory modality influences specific response properties of cortical neurons in another modality. Using loose-patch recording and Ca^{2+} imaging, we demonstrated that activation of auditory cortex sharpened OS of L2/3 excitatory neurons in V1, more so when visual contrast was low. The sharpening effect was attributed to an increase and decrease of response level at the preferred and orthogonal orientation respectively, with the overall response level reduced. These changes were achieved, at least partially, through interarea connectivity from A1 to V1: cortico-cortical axons from A1 L5 neurons mainly terminated in superficial layers of V1 and activated L1 inhibitory neurons. The latter likely caused the overall suppression of pyramidal cells in the underlying L2/3. In addition, L1 neurons could inhibit other inhibitory neurons in L2/3, generating a

disinhibitory effect that specifically contributed to the increased firing rate at the preferred orientation of the pyramidal cell. Different from OS, the same sound stimulation did not have a significant effect on direction selectivity (Figure S3F-S3G).

Auditory modulation of V1 processing through a top-down circuit

Our tracing study shows that in V1, the A1 axons project mostly to its superficial layers, with the highest axonal density observed in L1. This result is consistent with recent mouse brain connectome studies (Oh et al., 2014; Zingg et al., 2014). Such projection pattern is further supported by slice recording data: when A1 axons were optogenetically stimulated, L1 neurons (except those positive for VIP) received the maximum amount of direct input among all the major cell types in V1. In L2/3, pyramidal cells as well as PV, SOM and VIP inhibitory neurons also received direct A1 inputs but to a smaller, varying degree, with the inputs to pyramidal and PV neurons stronger than those to SOM and VIP cells. Neither type of neuron below 400 μm received direct A1 input. A previous study has reported a sound-induced hyperpolarization in V1 L2/3 pyramidal neurons, and it is suggested that a L5 to L2/3 inhibitory route underlies this suppressive effect (Iurilli et al., 2012). Our results do not support the latter notion. Instead, our data indicate that L1 neurons are the primary target of A1 input. Sound stimulation can directly excite V1 L1 neurons, which likely results in the hyperpolarization of pyramidal cells observed in Iurilli et al. (2012) study. Further supporting this notion are the observations that the onset latency of L1 neuron responses to sound alone (~ 40 ms, Figure 6G) was comparable to that of the hyperpolarization previously observed (Iurilli et al., 2012), and that spontaneous firing rates of L2/3 pyramidal cells are reduced when L1 is activated either by sound or by optogenetic stimulation of A1-V1 axons (Figure S1B). L1 neurons connect to all the cell types, including all inhibitory types, in the underlying L2/3 (Jiang et al., 2013; Lee et al., 2015; Wozny and Williams, 2011; Xu and Callaway, 2009; Zagha et al., 2013), and thereby are in a suitable position to provide general inhibition to L2/3 pyramidal cells and to produce disinhibitory effects by inhibiting certain L2/3 inhibitory neurons. Such L1-mediated top-down circuit is a likely candidate for mediating the auditory modulation of visual processing, since suppressing L1 activity reduces the sound effect of sharpening OS. It is thus tempting to postulate that L1 can serve as a hub of top-down modulation, receiving inputs from different sensory modalities or even non-sensory higher cortical areas (Zhang et al., 2014). Interestingly, while A1 profoundly innervates V1, viral injections in V1 in our experiments did not reveal any connectivity from V1 to A1 (Figure S4B, also see (Oh et al., 2014)). This suggests that direct connections between A1 and V1 may be unidirectional, at least in some species.

Sound effect on L1 neuron responses

Consistent with an involvement of L1, sound stimulation resulted in an elevation of visual responses in L1 neurons. The effect was dependent on the sound intensity. The minimum test intensity at which there was a significant increase in L1 responses was 60 dB SPL (Figure S6B), while the intensity threshold of A1 neurons was lower than 40 dB SPL (Li et al., 2014; Liu et al., 2007; Sun et al., 2013). The largest increase occurred when the maximum test intensity (80 dB SPL) was applied (Figure S6B). These results suggest that there is a certain intensity threshold for the cross-modal interaction coming into play. Interestingly, the relative increase of response level in L1 neurons was larger at their

preferred than orthogonal orientation (Figure 5F-5G), allowing them to produce larger disinhibitory effects at the preferred orientation. Such “supralinear” effect is possible and could be mediated by certain voltage-dependent conductances such as NMDA receptors and dendritic Ca^{2+} channels (Lavzin et al., 2012; Xu et al., 2012).

L2/3 Inhibitory neuron subtypes

PV, SOM and VIP neurons together account for nearly 100% of the inhibitory cell population, and VIP neurons comprise the largest inhibitory subgroup in L2/3 (Rudy et al., 2011). All of these inhibitory subtypes receive some direct A1 input, but to varying degrees. They are also innervated by L1 inhibitory neurons (Jiang et al., 2013). Additionally, within the L2/3 inhibitory neuron population, there is complex inhibition between inhibitory cells (Pfeffer et al., 2013). These interactions together predict that the effects of sound on L2/3 inhibitory neuron responses could be complex, which may depend on the delicate balance between the excitation from A1, inhibition from L1 of V1, and intra-laminar or local inhibition. Indeed, for PV neurons, we did not observe any significant changes of response levels in the presence of sound. PV neurons are thus unlikely to play a significant role in mediating the sound-induced effect on OS of pyramidal cells. The number of visually driven SOM neurons was so low that neither could these cells be a major player.

What types of inhibitory neuron in L2/3 are more likely to contribute to the sound modulation of OS, in particular the disinhibitory effect seen at the preferred orientation of pyramidal cells? Several previous studies have suggested that VIP neurons, the most abundant L2/3 inhibitory neurons, are the major driver of a disinhibitory circuit: their activation by various long-range projections inhibits SOM neurons, which in turn disinhibits pyramidal cells (Fu et al., 2014; Lee et al., 2013; Pfeffer et al., 2013; Pi et al., 2013). Contrary to what is expected from previous studies, VIP neurons were in fact suppressed in the presence of sound. There was a decrease in their overall Ca^{2+} response level, and an even greater decrease of the response at the preferred orientation. The latter effect may overcome the increased L1 inhibition in pyramidal cells, resulting in an overall disinhibitory effect at this stimulus orientation. Since excitation and inhibition in an excitatory cell display a similar preferred orientation (Li et al., 2012b; Liu et al., 2011), it is possible that the connected VIP and pyramidal cells share similar orientation preferences. Although a slice recording study shows that a VIP neuron's connection to a pyramidal cell is relatively weak (Pfeffer et al., 2013), optogenetic activation of VIP neurons causes a significant decrease of firing rate in at least a subset of pyramidal cells (Lee et al., 2012), indicating that VIP neurons as a population are able to modulate pyramidal cell firing rates. Together, the behavior of VIP neurons in the presence of sound, i.e. preferentially decreasing their firing rate at the preferred orientation, suggests that these inhibitory neurons are in a suitable position to mediate a disinhibitory effect at the preferred orientation of their connected pyramidal cells.

Altogether, our results demonstrate that auditory input can modulate visual processing in V1 principal neurons through a projection from A1 to V1 superficial layers, in particular L1. Such cross-modality top-down modulation may be abundant in sensory systems. It is likely advantageous for animals’ survival in that other senses (audition in this case, and probably

somatosensation also) can aid the visual system by improving the performance and discriminability of V1 principal neurons, especially under low visual contrast conditions. As such, barely detectable images of predators in a dim environment may be better perceived by rodents when there are sounds coincidentally associated with the threats. Arising from the current results are several intriguing questions: how this top-down circuit is formed during development; whether it is hardwired or if it can be shaped by visual or auditory experience (Petrus et al., 2014; Petrus et al., 2015). Studying the A1-V1 circuit in the developing brain of both naïve and experientially manipulated animals will thus be an exciting future research direction.

Experimental Procedures

All procedures were carried out in accordance with USC Animal Care and Use Committee. Adult female mice were used. The Cre lines used in this study were obtained from The Jackson Laboratory (with C57BL/6J background). Viral vectors, AAV2/9.EF1α.DIO.hChR2(H134R)-EYFP.WPRE.hGH, AAV9.hSyn.hChR2(H134R)-eYFP.WPRE.hGH, AAV2/1.Syn.GCaMP6s.WPRE.SV40, AAV2/1.Syn.Flex.GCaMP6s.WPRE.SV40, AAV1-CAG-FLEX-ArchT-GFP, were obtained from UPenn Vector Core and UNC vector core. Viral injections were performed as we previously reported (Liang et al., 2015; Xiong et al., 2015) (see Supplemental Experimental Procedures: Viral injections). Injections of retrograde tracers followed a previous study (Zingg et al., 2014) (see Supplemental Experimental Procedures: Retrograde tracer injection). For imaging brain slices, the animal was deeply anesthetized and transcardially perfused with 4% paraformaldehyde. Brain tissue was sliced into 150 μm sections using a vibratome (Leica), and sections were mounted onto glass slides and imaged under a confocal microscope (Olympus).

For *in vivo* recordings, we sedated the mouse with an intramuscular injection of chlorprothixene hydrochloride (10 mg/kg in 4 mg/ml water solution) and then anesthetized it with urethane (1.2 g/kg, i.p., at 20% w/v in saline). Surgery procedures followed previously studies (Li et al., 2013; Liu et al., 2009; Mangini and Pearlman, 1980; Niell and Stryker, 2008) (see Supplemental Experimental Procedures: Animal surgery). For awake recordings, one week before experiments, mice were anesthetized with isoflurane (1.5% by volume) and a screw for head fixation was mounted on top of the skull with dental cement. Surgery procedures followed our previous studies (Xiong et al., 2013; Zhou et al., 2014) (see Supplemental Experimental Procedures: Animal surgery). On the day of recording, the head screw was tightly fit into a metal post and the mouse was allowed to run smoothly on a rotatable plate (Liang et al., 2015). Due to potential complex effects of locomotion on visual responses (Fu et al., 2014; Niell and Stryker, 2010), we have only selected recording trials in which animals remained stationary.

Cell-attached loose-patch recordings were performed with an Axopatch 200B amplifier (Molecular Devices) using 4-6 MΩ glass pipettes, following our previous studies (Liang et al., 2015; Liu et al., 2009) (see Supplemental Experimental Procedures: *In vivo* electrophysiology). The spike signal was filtered at 10 kHz and sampled at 20 kHz. All the neurons recorded under this condition showed regular-spikes (the spike waveform had a

trough-to-peak interval of 0.85 ± 0.10 ms, n = 72 cells), consistent with the sampling bias towards excitatory neurons as shown previously with cell morphology reconstructions (Liu et al., 2010; Wu et al., 2008). The recordings were done in L2/3 (200-350 μm from the pia) and L4 (375-510 μm). The layer assignment of the blindly recorded neurons was made mostly according to the vertical travel distance of the electrode. To record from fast-spiking (FS) inhibitory neurons, smaller pipettes with a higher impedance (10 MΩ) were used and neurons with fast-spikes (trough-to-peak interval of the spike waveform < 0.5 ms) were actively searched (Wu et al., 2008). We made recordings in the monocular zone of the V1.

To measure orientation tuning, we applied drifting sinusoidal gratings (spatial frequency of 0.04 cycles per degree and temporal frequency of 2Hz) of 12 directions (30° steps) in a random sequence. We recorded spontaneous activity when presenting a uniform grey background. The visual stimulation with and without sound or LED illumination were alternated, but the stimulus sequence was randomized independently for sound/LED off and sound/LED on trials. We set the inter-stimulus interval at 10 s to allow a full recovery of ChR2 function from desensitization (Li et al., 2013). Each cell was recorded under high contrast (95%) and low contrast (25%) conditions. We applied five to ten sets of stimuli to each cell, with the sequence different between sets. The long recording time in our experimental conditions prevented us from applying different combinations of spatial and temporal frequency. To better drive as large fraction as possible of V1 neurons, we have carefully chosen our visual stimuli (0.04 cpd, 2Hz) based on a previous study of a large number of neurons (Niell and Stryker, 2008), which shows that the largest fraction of mouse V1 cells prefers the spatial frequency of 0.04 cpd, and the average preferred spatial frequency in L2/3 is ~0.04 cpd, and that most of units prefer a temporal frequency of ~2Hz. Since the level of change in selectivity was not correlated with the initial selectivity level or the overall response level (Supplementary Figure 4D-4F), it is likely that both optimally driven and non-optimally driven neurons could increase their selectivity under sound/A1 stimulation.

The auditory stimulation consisted of white noise pulses at 70 dB sound pressure level (SPL) presented at 10Hz (50 ms on, 50 ms off) throughout the duration of the visual stimulus. The onset and offset of auditory stimulation were the same as visual stimulation. The sound was delivered by a single speaker located contralateral to the recording side. The visual stimuli alone and those coupled with sound were alternated, but the stimulus sequence was randomized independently for sound-off and sound-on trials.

To photoactivate hChR2, we used a blue (470 nm) fiber-coupled LED (0.8 mm diameter, Doric Lenses) placed on top of the exposed cortical surface. LED light spanned the entire area of V1. We applied black pigment stained agar to cover the tip of the optic fiber, as to prevent LED light leakage reaching the contralateral eye. We had verified that LED light did not directly stimulate the contralateral eye in wild-type mice (data not shown). The LED was driven by the analog output from a NIDAQ board (National Instruments). The intensity of LED was around 5 mW (measured at the tip of the fiber). To photoactivate ArchT, we used a green light (530 nm) fiber-coupled LED (0.8 mm diameter, Doric Lenses) and followed the same procedure as with hChR2.

Slice recordings were performed following our previous studies (Li et al., 2014; Ji et al., 2015) (see Supplemental Experimental Procedures: *In vitro* electrophysiology). Coronal cortical slices of 350 μm thickness were cut from the infected brain hemisphere by a vibrating microtome (Leica VT1000s). Recording was made under an upright fluorescence microscope (Olympus BX51WI) equipped with an infrared light source. Whole-cell voltage-clamp recordings were selectively performed on fluorescence-labeled inhibitory neurons in PV-Cre, SOM-Cre, VIP-Cre or GAD2-Cre::tdTomato slices or non-fluorescent excitatory cells in GAD2-Cre::tdTomato slices under epifluorescence imaging in V1. TTX (a sodium channel blocker, 1μM) and 4-aminopyridine (a potassium channel blocker, 1mM) were present in the external solution. Excitatory synaptic currents were recorded by clamping the cell's membrane potential at −70 mV. Signals were filtered at 2 kHz and sampled at 10 kHz. In each slice, multiple neurons were recorded. The cortical depth of each recorded cell was based on the vertical distance of the cell body from the pial surface of the cortex, which was set as 0 μm. The distance was measured with a micromanipulator coupled with a digital reader (SD Instrument DR1000). L1 was defined as 0 −150 μm. Blue light pulses were delivered from a mercury Arc lamp gated by an electronic shutter. The power was 3 mW measured at the focal plane. Brief light pulses (3 ms) were applied individually (0.033 Hz). For each condition, 10 – 30 trials were given and responses were averaged.

Ca^{2+} Imaging was performed after at least 2 weeks of viral expression. We used a custom built Mai Tai (Spectra-physics) based 2-photon system and recorded data using a custom-modified version of the Scanimage software (Pologruto et al. 2003). Imaging from labeled cells was performed in multiple subregions each spanning ~200 × 200 up to ~400 × 400 μm. Scan lines were drawn across each of the clearly visible cell bodies (typically 5-15 cells) and then imaged continuously (rapidly alternating between all the scan lines) while presenting moving sinusoidal grating stimuli (1.5 sec) at 10 second intervals. Depending on the number of scan lines the scan rate ranged between 25-70 Hz. Within each scan line, we manually defined regions of interests (ROIs) based on the presence of fluorescence transients. The selected ROIs were compared with the 2 dimensional snapshot of the region to make sure that we were imaging the labeled cell bodies rather than neuropils. The signal within a ROI was processed using standard methods to derive fractional change over baseline, i.e. ΔF/F0 (Jia et al. 2011). Neurons located within 150 μm below the pia surface were considered as L1 cells.

For analysis, we counted the spikes evoked by drifting sinusoidal gratings within a time window covering the visual stimulation duration with a 70 ms delay, and subtracted the average spontaneous firing rate from the stimulus-evoked spike rate. We quantified the strength of OS with a global orientation selectivity index (gOSI), which considers responses at all test orientations (Mariño et al., 2005):

$$gOSI = \| \sum R(\theta) \times e^{2i\theta} \| \sum R(\theta)$$

i is $\sqrt{-1}$. θ is the angle of the moving direction. $R(\theta)$ is the response level at angle θ. The preferred orientation was determined from this vector sum of all the responses. We then averaged responses to the gratings of opposite directions, and obtained the orientation tuning

curve between 0 – 180 degrees, and fitted it with a Gaussian function: $R(\theta) = A \times exp\left(-0.5 \times (\theta - \phi)^2/\sigma^2\right) + B$. φ is the determined preferred orientation and σ is the tuning width. We also computed an orientation selectivity index (OSI) defined as $(R_{pref} - R_{orth})/(R_{pref} + R_{orth})$, where R_{pref} is the response level at the angle of φ, and R_{orth} is that at the angle of $\varphi + 90$. The gOSI and OSI values vary between 0 and 1, with 0 being the value for an untuned neuron and 1 for a perfectly tuned neuron (Carandini and Ferster, 2000; Mariño et al., 2005; Ringach et al., 2002). To compare tuning between conditions, the preferred orientation was the peak of the Gaussian fit for the data in the sound/LED off condition (e.g. in Figure 2E, right panel), or was the test orientation closest to the peak of the Gaussian fit in the sound/LED off condition (e.g. in Figure 1B, 1C, 2E left panel, and 2G). Similarly, the orthogonal orientation (e.g. in Figure 1D, 2H) was also determined based on the data in the sound/LED off condition of each cell. Tuning curves were aligned according the preferred orientation (set as zero degree) before averaging. The global direction selectivity index (gDSI) was quantified according to:

$$gDSI = \| \sum R(\theta) \times e^{i\theta} \| \sum R(\theta)$$

For slice recording data, peak response (EPSC) amplitude was measured for each cell. In order to take into account the varying levels of channelrhodopsin expression among different slices, we quantified a relative response amplitude. Specifically, in each slice, at least three L2/3 pyramidal cells were recorded, and their EPSC amplitudes were averaged to obtain a mean pyramidal cell response for that slice. The mean pyramidal cell responses of different slices were then averaged to obtain a global average pyramidal cell response. In each slice, a scaling factor was determined by the ratio between the mean pyramidal cell response for that slice and the global average pyramidal cell response. EPSC amplitudes of all of cells recorded in the same slice were then scaled according to the scaling factor for that slice.

For statistical analysis, we first performed Lilliefors test to check whether the data were normally distributed. In the case of a normal distribution, we performed paired t-test. Otherwise, we performed a non-parametric test (Wilcoxon signed-rank test in this study). For the L1 inhibition experiment, one-way ANOVA with repeated measures with Bonferroni post-hoc test was performed. For the slice recording experiment, the relative input strengths of different groups were compared using one-way ANOVA with Bonferroni post-hoc test. No statistical method was used to pre-determine sample sizes, but our sample sizes were similar to those reported in previous publications in the field.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants from the US National Institutes of Health (EY019049 and EY025722 to H.W.T., DC008983 to L.I.Z.) and the Kirchgessner Foundation (to H.W.T). We thank Dr. X.R.Xiong for the technical help.

References

Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M. A neural circuit for spatial summation in visual cortex. Nature. 2012; 490:226–231. [PubMed: 23060193]

Andersen RA, Snyder RL, Merzenich MM. The topographic organization of corticocollicular projections from physiologically identified loci in the AI, AII, and anterior auditory cortical fields of the cat. J Comp Neurol. 1980; 191:479–494. [PubMed: 7410603]

Atallah BV, Bruns W, Carandini M, Scanziani M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron. 2012; 73:159–170. [PubMed: 22243754]

Bajo VM, Moore DR. Descending projections from the auditory cortex to the inferior colliculus in the gerbil, Meriones unguiculatus. J Comp Neurol. 2005; 486:101–116. [PubMed: 15844210]

Bajo VM, Nodal FR, Bizley JK, Moore DR, King AJ. The ferret auditory cortex: descending projections to the inferior colliculus. Cereb Cortex. 2007; 17:475–491. [PubMed: 16581982]

Beaulieu C, Kisvarday Z, Somogyi P, Cynader M, Cowey A. Quantitative distribution of GABA-immunopositive and -immunonegative neurons and synapses in the monkey striate cortex (area 17). Cereb Cortex. 1992; 2:295–309. [PubMed: 1330121]

Campi KL, Bales KL, Grunewald R, Krubitzer L. Connections of auditory and visual cortex in the prairie vole (Microtus ochrogaster): evidence for multisensory processing in primary sensory areas. Cereb Cortex. 2010; 20:89–108. [PubMed: 19395525]

Cappe C, Barone P. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur J Neurosci. 2005; 22:2886–2902. [PubMed: 16324124]

Carandini M, Ferster D. Membrane potential and firing rate in cat primary visual cortex. J Neurosci. 2000; 20:470–484. [PubMed: 10627623]

Chu Z, Galarreta M, Hestrin S. Synaptic interactions of late-spiking neocortical neurons in layer 1. J Neurosci. 2003; 23:96–102. [PubMed: 12514205]

Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. Cogn Affect Behav Neurosci. 2004; 4:117–126. [PubMed: 15460918]

Cruikshank SJ, Urabe H, Nurmikko AV, Connors BW. Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. Neuron. 2010; 65:230–245. [PubMed: 20152129]

Dewson JH, Cowey A, Weiskrantz L. Disruptions of auditory sequence discrimination by unilateral and bilateral cortical ablations of superior temporal gyrus in the monkey. Exp Neurol. 1970; 28:529–548. [PubMed: 4990931]

Driver J, Spence C. Cross-modal links in spatial attention. Philos Trans R Soc Lond B Biol Sci. 1998; 353:1319–1331. [PubMed: 9770225]

Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci. 2002; 22:5749–5759. [PubMed: 12097528]

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex. 1991; 1:1–47. [PubMed: 1822724]

Frens MA, Van Opstal AJ, Van der Willigen RF. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. Percept Psychophys. 1995; 57:802–816. [PubMed: 7651805]

Frostig RD, Xiong Y, Chen-Bee CH, Kvasnák E, Stehberg J. Large-scale organization of rat sensorimotor cortex based on a motif of large activation spreads. J Neurosci. 2008; 28:13274–13284. [PubMed: 19052219]

Fu Y, Tucciarone JM, Espinosa JS, Sheng N, Darcy DP, Nicoll RA, Huang ZJ, Stryker MP. A cortical circuit for gain control by behavioral state. Cell. 2014; 156:1139–1152. [PubMed: 24630718]

Gielen SC, Schmidt RA, Van den Heuvel PJ. On the nature of intersensory facilitation of reaction time. Percept Psychophys. 1983; 34:161–168. [PubMed: 6634374]

Gleiss S, Kayser C. Audio-visual detection benefits in the rat. PLoS One. 2012; 7:e45677. [PubMed: 23029179]

HERSHENSON M. Reaction time as a measure of intersensory facilitation. J Exp Psychol. 1962; 63:289–293. [PubMed: 13906889]

Iurilli G, Ghezzi D, Olcese U, Lassi G, Nazzaro C, Tonini R, Tucci V, Benfenati F, Medini P. Sound-driven synaptic inhibition in primary visual cortex. Neuron. 2012; 73:814–828. [PubMed: 22365553]

Jaekl PM, Harris LR. Sounds can affect visual perception mediated primarily by the parvocellular pathway. Vis Neurosci. 2009; 26:477–486. [PubMed: 20137106]

Ji XY, Zingg B, Mesik L, Xiao Z, Zhang LI, Tao HW. Thalamocortical Innervation Pattern in Mouse Auditory and Visual Cortex: Laminar and Cell-Type Specificity. Cereb Cortex. 2015

Jiang X, Wang G, Lee AJ, Stornetta RL, Zhu JJ. The organization of two new cortical interneuronal circuits. Nat Neurosci. 2013; 16:210–218. [PubMed: 23313910]

Jones EG, Powell TP. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain. 1970; 93:793–820. [PubMed: 4992433]

Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. Neuron. 2005; 48:373–384. [PubMed: 16242415]

Kok P, Jehee JF, de Lange FP. Less is more: expectation sharpens representations in the primary visual cortex. Neuron. 2012; 75:265–270. [PubMed: 22841311]

KUYPERS HG, SZWARCBART MK, MISHKIN M, ROSVOLD HE. OCCIPITOTEMPORAL CORTICOCORTICAL CONNECTIONS IN THE RHESUS MONKEY. Exp Neurol. 1965; 11:245–262. [PubMed: 14295621]

Lavzin M, Rapoport S, Polsky A, Garion L, Schiller J. Nonlinear dendritic processing determines angular tuning of barrel cortex neurons in vivo. Nature. 2012; 490:397–401. [PubMed: 22940864]

Lee AJ, Wang G, Jiang X, Johnson SM, Hoang ET, Lanté F, Stornetta RL, Beenhakker MP, Shen Y, Julius Zhu J. Canonical Organization of Layer 1 Neuron-Led Cortical Inhibitory and Disinhibitory Interneuronal Circuits. Cereb Cortex. 2015; 25:2114–2126. [PubMed: 24554728]

Lee S, Kruglikov I, Huang ZJ, Fishell G, Rudy B. A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nat Neurosci. 2013; 16:1662–1670. [PubMed: 24097044]

Lee SH, Kwan AC, Zhang S, Phoumthipphavong V, Flannery JG, Masmanidis SC, Taniguchi H, Huang ZJ, Zhang F, Boyden ES, et al. Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature. 2012; 488:379–383. [PubMed: 22878719]

Li J, Schwark HD. Distribution and proportions of GABA-immunoreactive neurons in cat primary somatosensory cortex. J Comp Neurol. 1994; 343:353–361. [PubMed: 7517965]

Li LY, Ji XY, Liang F, Li YT, Xiao Z, Tao HW, Zhang LI. A feedforward inhibitory circuit mediates lateral refinement of sensory representation in upper layer 2/3 of mouse primary auditory cortex. J Neurosci. 2014; 34:13670–13683. [PubMed: 25297094]

Li LY, Xiong XR, Ibrahim LA, Yuan W, Tao HW, Zhang LI. Differential Receptive Field Properties of Parvalbumin and Somatostatin Inhibitory Neurons in Mouse Auditory Cortex. Cereb Cortex. 2015a; 25:1782–1791. [PubMed: 24425250]

Li YT, Ibrahim LA, Liu BH, Zhang LI, Tao HW. Linear transformation of thalamocortical input by intracortical excitation. Nat Neurosci. 2013; 16:1324–1330. [PubMed: 23933750]

Li YT, Liu BH, Chou XL, Zhang LI, Tao HW. Synaptic Basis for Differential Orientation Selectivity between Complex and Simple Cells in Mouse Visual Cortex. J Neurosci. 2015b; 35:11081–11093. [PubMed: 26245969]

Li YT, Ma WP, Li LY, Ibrahim LA, Wang SZ, Tao HW. Broadening of inhibitory tuning underlies contrast-dependent sharpening of orientation selectivity in mouse visual cortex. J Neurosci. 2012a; 32:16466–16477. [PubMed: 23152629]

Li YT, Ma WP, Pan CJ, Zhang LI, Tao HW. Broadening of cortical inhibition mediates developmental sharpening of orientation selectivity. J Neurosci. 2012b; 32:3981–3991. [PubMed: 22442065]

Liang F, Xiong XR, Zingg B, Ji XY, Zhang LI, Tao HW. Sensory Cortical Control of a Visually Induced Arrest Behavior via Corticotectal Projections. Neuron. 2015; 86:755–767. [PubMed: 25913860]

Liang M, Mouraux A, Hu L, Iannetti GD. Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. Nat Commun. 2013; 4:1979. [PubMed: 23752667]

Liu BH, Li P, Li YT, Sun YJ, Yanagawa Y, Obata K, Zhang LI, Tao HW. Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. J Neurosci. 2009; 29:10520–10532. [PubMed: 19710305]

Liu BH, Li P, Sun YJ, Li YT, Zhang LI, Tao HW. Intervening inhibition underlies simple-cell receptive field structure in visual cortex. Nat Neurosci. 2010; 13:89–96. [PubMed: 19946318]

Liu BH, Li YT, Ma WP, Pan CJ, Zhang LI, Tao HW. Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells. Neuron. 2011; 71:542–554. [PubMed: 21835349]

Liu BH, Wu GK, Arbuckle R, Tao HW, Zhang LI. Defining cortical frequency tuning with recurrent excitatory circuitry. Nat Neurosci. 2007; 10:1594–1600. [PubMed: 17994013]

Ma WP, Liu BH, Li YT, Huang ZJ, Zhang LI, Tao HW. Visual representations by cortical somatostatin inhibitory neurons--selective but with weak and delayed responses. J Neurosci. 2010; 30:14371–14379. [PubMed: 20980594]

Mangini NJ, Pearlman AL. Laminar distribution of receptive field properties in the primary visual cortex of the mouse. J Comp Neurol. 1980; 193:203–222. [PubMed: 6776165]

Mariño J, Schummers J, Lyon DC, Schwabe L, Beck O, Wiesing P, Obermayer K, Sur M. Invariant computations in local cortical networks with balanced excitation and inhibition. Nat Neurosci. 2005; 8:194–201. [PubMed: 15665876]

Markovitz CD, Tang TT, Lim HH. Tonotopic and localized pathways from primary auditory cortex to the central nucleus of the inferior colliculus. Front Neural Circuits. 2013; 7:77. [PubMed: 23641201]

McDonald JJ, Teder-Sälejärvi WA, Hillyard SA. Involuntary orienting to sound improves visual perception. Nature. 2000; 407:906–908. [PubMed: 11057669]

Niell CM, Stryker MP. Highly selective receptive fields in mouse visual cortex. J Neurosci. 2008; 28:7520–7536. [PubMed: 18650330]

Niell CM, Stryker MP. Modulation of visual responses by behavioral state in mouse visual cortex. Neuron. 2010; 65:472–479. [PubMed: 20188652]

Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, et al. A mesoscale connectome of the mouse brain. Nature. 2014; 508:207–214. [PubMed: 24695228]

Petreanu L, Mao T, Sternson SM, Svoboda K. The subcellular organization of neocortical excitatory connections. Nature. 2009; 457:1142–1145. [PubMed: 19151697]

Petrus E, Isaiah A, Jones AP, Li D, Wang H, Lee HK, Kanold PO. Crossmodal induction of thalamocortical potentiation leads to enhanced information processing in the auditory cortex. Neuron. 2014; 81:664–673. [PubMed: 24507197]

Petrus E, Rodriguez G, Patterson R, Connor B, Kanold PO, Lee HK. Vision Loss Shifts the Balance of Feedforward and Intracortical Circuits in Opposite Directions in Mouse Primary Auditory and Visual Cortices. J Neurosci. 2015; 35:8790–8801. [PubMed: 26063913]

Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat Neurosci. 2013; 16:1068–1076. [PubMed: 23817549]

Pi HJ, Hangya B, Kvitsiani D, Sanders JI, Huang ZJ, Kepecs A. Cortical interneurons that specialize in disinhibitory control. Nature. 2013; 503:521–524. [PubMed: 24097352]

Ringach DL, Shapley RM, Hawken MJ. Orientation selectivity in macaque V1: diversity and laminar dependence. J Neurosci. 2002; 22:5639–5651. [PubMed: 12097515]

Rudy B, Fishell G, Lee S, Hjerling-Leffler J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol. 2011; 71:45–61. [PubMed: 21154909]

Sieben K, Röder B, Hanganu-Opatz IL. Oscillatory entrainment of primary somatosensory cortex encodes visual control of tactile processing. J Neurosci. 2013; 33:5736–5749. [PubMed: 23536087]

Stehberg J, Dang PT, Frostig RD. Unimodal primary sensory cortices are directly connected by long-range horizontal projections in the rat sensory cortex. Front Neuroanat. 2014; 8:93. [PubMed: 25309339]

Sun YJ, Kim YJ, Ibrahim LA, Tao HW, Zhang LI. Synaptic mechanisms underlying functional dichotomy between intrinsic-bursting and regular-spiking neurons in auditory cortical layer 5. J Neurosci. 2013; 33:5326–5339. [PubMed: 23516297]

Taniguchi H, He M, Wu P, Kim S, Paik R, Sugino K, Kvitsiani D, Kvitsani D, Fu Y, Lu J, et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron. 2011; 71:995–1013. [PubMed: 21943598]

Vroomen J, de Gelder B. Sound enhances visual perception: cross-modal effects of auditory organization on vision. J Exp Psychol Hum Percept Perform. 2000; 26:1583–1590. [PubMed: 11039486]

Winans SS. Visual form discrimination after removal of the visual cortex in cats. Science. 1967; 158:944–946. [PubMed: 6054170]

Winer JA, Larue DT. Populations of GABAergic neurons and axons in layer I of rat auditory cortex. Neuroscience. 1989; 33:499–515. [PubMed: 2636704]

Wozny C, Williams SR. Specificity of synaptic connectivity between layer 1 inhibitory interneurons and layer 2/3 pyramidal neurons in the rat neocortex. Cereb Cortex. 2011; 21:1818–1826. [PubMed: 21220765]

Wu GK, Arbuckle R, Liu BH, Tao HW, Zhang LI. Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. Neuron. 2008; 58:132–143. [PubMed: 18400169]

Xiong XR, Liang F, Zingg B, Ji XY, Ibrahim LA, Tao HW, Zhang LI. Auditory cortex controls sound-driven innate defense behaviour through corticofugal projections to inferior colliculus. Nat Commun. 2015; 6:7224. [PubMed: 26068082]

Xu NL, Harnett MT, Williams SR, Huber D, O'Connor DH, Svoboda K, Magee JC. Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature. 2012; 492:247–251. [PubMed: 23143335]

Xu X, Callaway EM. Laminar specificity of functional input to distinct types of inhibitory cortical neurons. J Neurosci. 2009; 29:70–85. [PubMed: 19129386]

Zagha E, Casale AE, Sachdev RN, McGinley MJ, McCormick DA. Motor cortex feedback influences sensory processing by modulating network state. Neuron. 2013; 79:567–578. [PubMed: 23850595]

Zhang S, Xu M, Kamigaki T, Hoang Do JP, Chang WC, Jenvay S, Miyamichi K, Luo L, Dan Y. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science. 2014; 345:660–665. [PubMed: 25104383]

Zhou FM, Hablitz JJ. Morphological properties of intracellularly labeled layer I neurons in rat neocortex. J Comp Neurol. 1996; 376:198–213. [PubMed: 8951637]

Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, Tao HW, Zhang LI. Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. Nat Neurosci. 2014; 17:841–850. [PubMed: 24747575]

Zingg B, Hintiryan H, Gou L, Song MY, Bay M, Bienkowski MS, Foster NN, Yamashita S, Bowman I, Toga AW, Dong HW. Neural networks of the mouse neocortex. Cell. 2014; 156:1096–1111. [PubMed: 24581503]

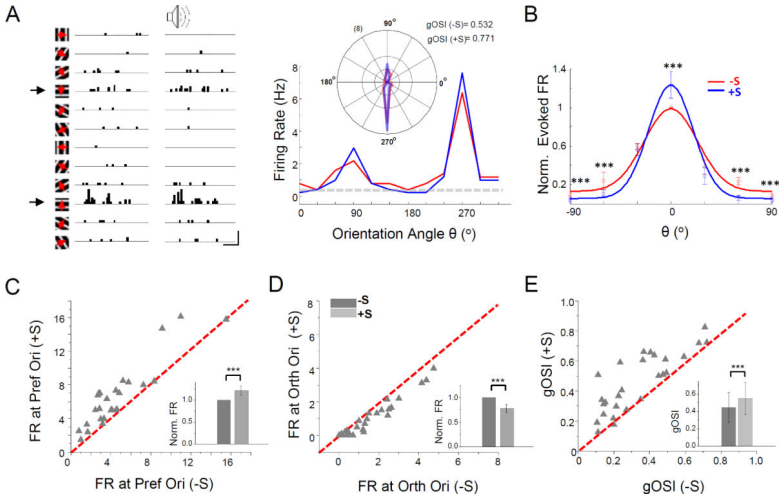

Figure 1. Sound induced sharpening of orientation selectivity in L2/3 of V1

(A) Left, post-stimulus spike time histograms (PSTHs) of spike responses of an example L2/3 pyramidal neuron to drifting sinusoidal gratings without (left) and with (right) coupling with sound stimulation, examined with *in vivo* loose-patch recording. Arrows mark the preferred orientation of the cell. Scale: 60 Hz, 500 ms. Right, firing rates at different stimulus directions for the same cell. Red: visual stimulation only. Blue: visual plus sound stimulation. Dashed grey line marks the average spontaneous firing rate. Inset, polar graph plotting. The axial value (firing rate) is indicated within the parentheses.

(B) Average normalized firing rates for 26 L2/3 pyramidal cells without (red) and with (blue) sound, fitted with a Gaussian function. The tuning curves were aligned according to the preferred orientation (designated as 0°) in the sound-off condition. Responses to the two opposite directions were averaged. Error bar = SEM. ***, $p < 0.001$, Wilcoxon signed-rank test.

(C-E) Plots of evoked firing rate at the preferred (C) and orthogonal orientation (D), as well as of gOSI value (E), in the sound-on (+S) versus sound-off (−S) condition for individual cells. The preferred and orthogonal orientations represent are those determined for the cell in the sound-off condition. Insets, mean evoked firing rates (normalized to the value in the sound-off condition) (C, D), or mean gOSI values (E). ***, $p < 0.001$, Wilcoxon signed-rank test (C) or paired t-test (D, E). Error bar = SEM (C, D) or SD (E).

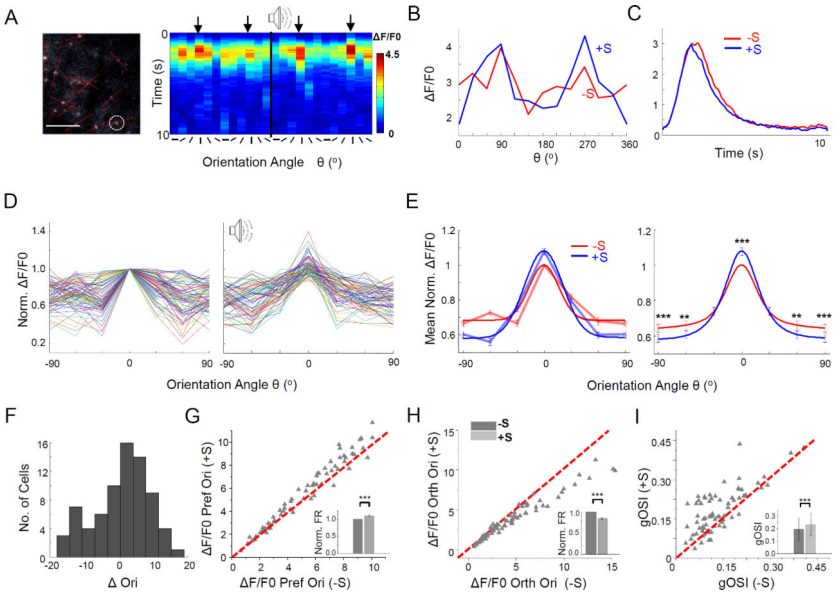

Figure 2. Sound-induced sharpening of OS revealed by two-photon Ca^{2+} imaging

(A) Left, an example image plane in L2/3 (250 μm below the pia). Red lines were used to guide line scanning across labeled cell bodies. Scale bar: 100 μm. Right, color map of ΔF/F0 (average of 5 repetitions) over an imaging period of 10 sec for 12 stimulus directions, plotted for an example cell (marked with a white circle on the left image). Time zero represents the onset of visual stimuli. Orientation is indicated by a black line below the color map. The first block shows visual responses alone; the second shows the responses to visual plus sound stimulation. Black arrows point to the preferred orientation.

(B) Tuning curves of peak ΔF/F0 for the cell shown in (A).

(C) Average Ca^{2+} responses across all orientations for the same cell.

(D) Tuning curves of normalized peak Ca^{2+} response amplitude plotted for 75 imaged L2/3 cells in the absence (left) and presence (right) of sound. Tuning curves were aligned according to the preferred orientation determined in the sound-off condition.

(E) Left, average normalized Ca^{2+} response levels of all the cells. Bright red and blue curves represent the Gaussian fits. Right, average of Gaussian fits of individual cell tuning curves. **, $p < 0.01$; ***, $p < 0.001$, Wilcoxon signed-rank test.

(F) Histogram of differences in preferred orientation between sound-off and sound-on conditions for cells included in analysis. Preferred orientations were determined from Gaussian fits.

(G-I) Plots of ΔF/F0 at the preferred (G) and orthogonal orientation (H), as well as of gOSI value (I) in the sound-on versus sound-off condition (n = 75 cells). Insets, mean normalized evoked Ca^{2+} response levels (G, H), or mean gOSI values (I). ***, $p < 0.001$, Wilcoxon signed-rank test. Error bar = SEM (G, H) or SD (I).

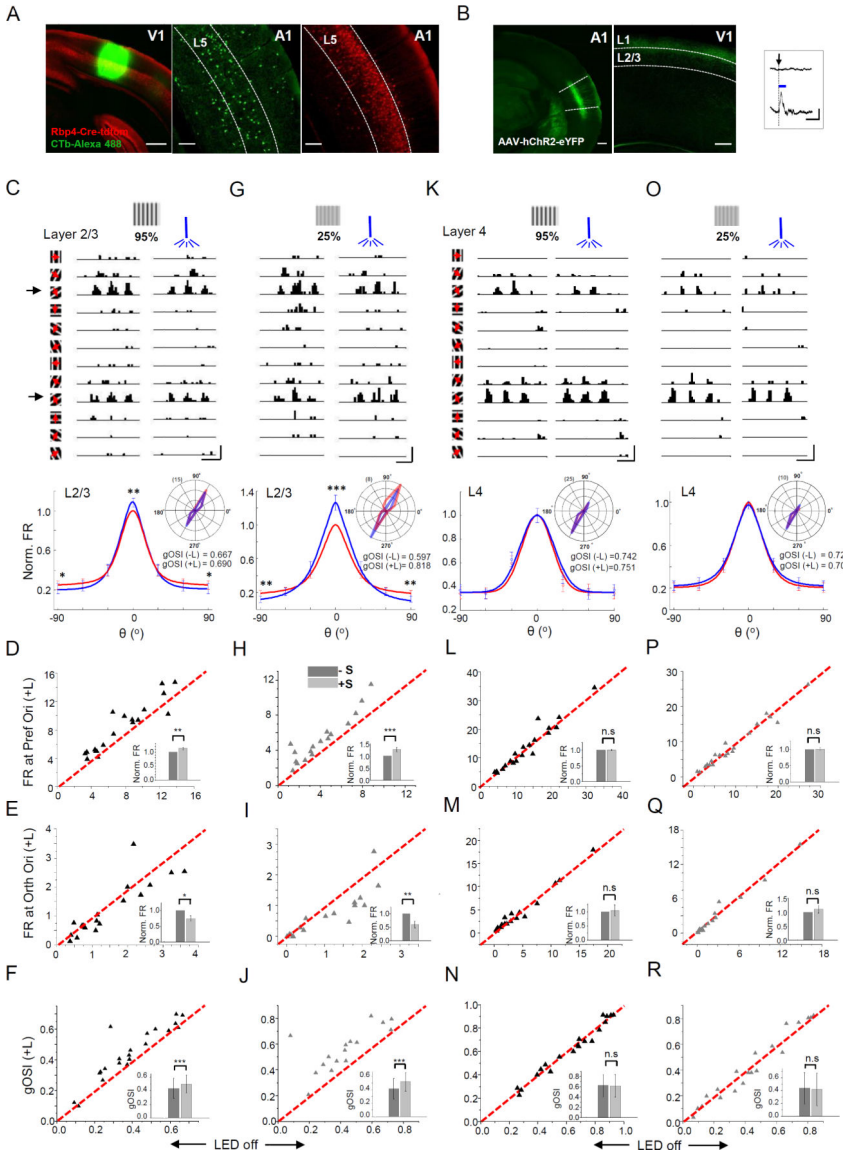

Figure 3. Optogenetic activation of A1 axons in V1

(A) Retrograde labeling. Left, CTb-Alexa 488 (green) injection site in V1 of a Rbp4-Cre::tdTomato (red) mouse. Middle, retrogradely labeled neurons in A1. Right, tdTomato expression pattern in A1. Scale bars, 200 μm.

(B) Anterograde labeling. Left, AAV-DIO-hChR2-eYFP injection site in A1 (outlined by two dash lines) of a Rbp4-Cre mouse. Scale bar, 250 μm. Middle, A1 axons in V1 of the same animal. Scale bar, 200 μm. Right, example local field currents recorded in superficial layers of V1 without (upper, trigger only) and with (lower) LED illumination (duration = 50 ms, onset marked by the arrow). Scale: 20 pA, 100 ms.

(C) Top, PSTHs for spike responses of an example L2/3 pyramidal neuron to gratings at a high (95%) contrast, without (left) and with (right) LED illumination. Scale: 70 Hz, 500 ms. Bottom, average normalized evoked firing rate (FR) at different orientations (relative to the preferred) for all the recorded L2/3 cells. Inset, polar graph plotting (inset) for the cell shown in top. Red, visual stimulation only; blue, visual plus LED stimulation.

(D-F) Plots of evoked firing rate at the preferred (D) and orthogonal (E) orientation, as well as of gOSI (F) in the LED-on versus LED-off conditions for all the cells (n = 18). Insets, mean normalized evoked firing rates (D, E) and mean gOSI values (F) of all the cells. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, paired t-test (D, E) or Wilcoxon signed-rank test (F). Error bar = SEM (D, E) or SD (F).

(G) The same cell as in (C), except that visual stimuli were at a low (25%) contrast. Scale: 60 Hz, 500 ms.

(H-J) Responses under low contrast stimuli. Data are presented in similar manners as in (D-F) (n = 18). **, $p < 0.01$; ***, $p < 0.001$, Wilcoxon signed-rank test (H) or paired t-test (I, J).

(K-N) Responses of L4 neurons to high contrast stimuli (n = 20 cells). Data are presented in similar manners as in (C-F). Scale in (K): 80 Hz, 500 ms. “n.s”, nonsignificant, $p > 0.05$, Wilcoxon signed-rank test (L, M) or paired t-test (N).

(O-R) Responses of L4 neurons to low contrast visual stimuli (n = 20 cells). Scale in (O): 80 Hz, 500 ms.

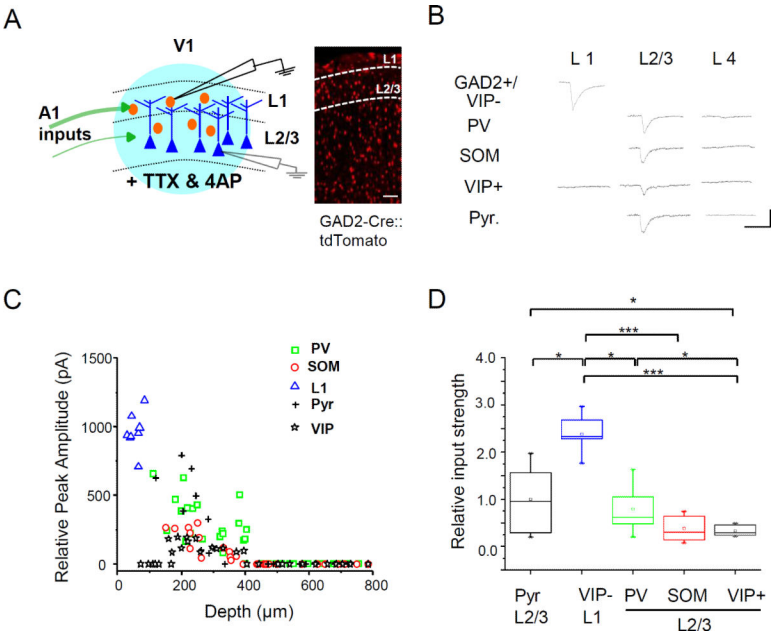

Figure 4. L1 neurons receive maximum A1 input

(A) Left, schematic diagram of whole-cell recordings in a slice preparation. The thickness of green arrows represents density of A1 axons in each layer. Red, inhibitory neuron; dark blue, pyramidal neuron; light blue, LED illumination. Right, fluorescence-labeled inhibitory neurons in V1 of a GAD2-Cre::tdTomato mouse. Scale bar, 200 μm.

(B) Average excitatory postsynaptic current (EPSC) traces of example neurons of different types in different layers evoked by LED stimulation. Note that in L1, VIP+ cells did not show EPSCs. Scale: 50 pA, 50 ms.

(C) Relative peak EPSC amplitudes of all the recorded cells versus their depths. Blue marks VIP-negative L1 neurons.

(D) Summary of relative strength of A1 input to different types of neurons (n = 12 for pyramidal, 9 for VIP- L1, 19 for PV, 12 for SOM, 10 for VIP+ L2/3). Small box, mean value; large box, SD; whisker, full range; line, median. *, $p < 0.05$; ***, $p < 0.001$, one-way ANOVA and post hoc test.

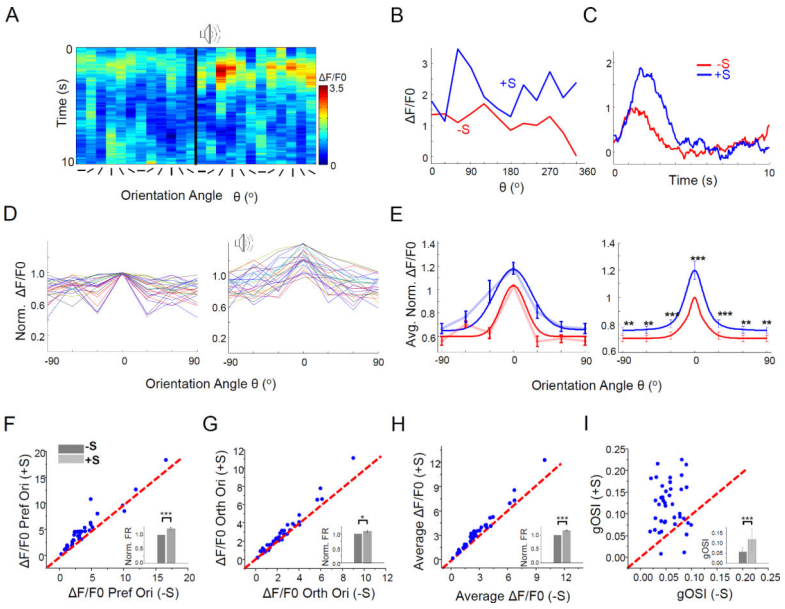

Figure 5. Visual responses of L1 neurons are enhanced by sound

(A) Color map of Ca^{2+} responses of an example L1 neuron to gratings (average of 5 repetitions) in the absence (left block) and presence (right) of sound stimulation.

(B) Tuning curves of peak Ca^{2+} response level in the absence (red) and presence (blue) of sound, plotted for the same cell shown in (A).

(C) Time-dependent ΔF/F0 averaged across all orientations for the same cell.

(D) Normalized tuning curves in the absence (left) and presence (right) of sound for all imaged L1 neurons (n = 40).

(E) Left, average normalized Ca^{2+} response amplitudes of all L1 neurons. Bright red and blue curves are Gaussian fits. Right, average of all Gaussian fits of individual L1 cells.

(F-I) Plots of peak Ca^{2+} response amplitude at the preferred (F) and orthogonal orientation (G), of average response amplitude across all orientations (H), as well as of gOSI value (I) in the sound-on versus sound-off conditions for all L1 cells. Insets, average values. *, $p < 0.05$; ***, $p < 0.001$, Wilcoxon signed-rank test (E, F) or paired t-test (G, H). Error bar = SEM (E-G) or SD (H).

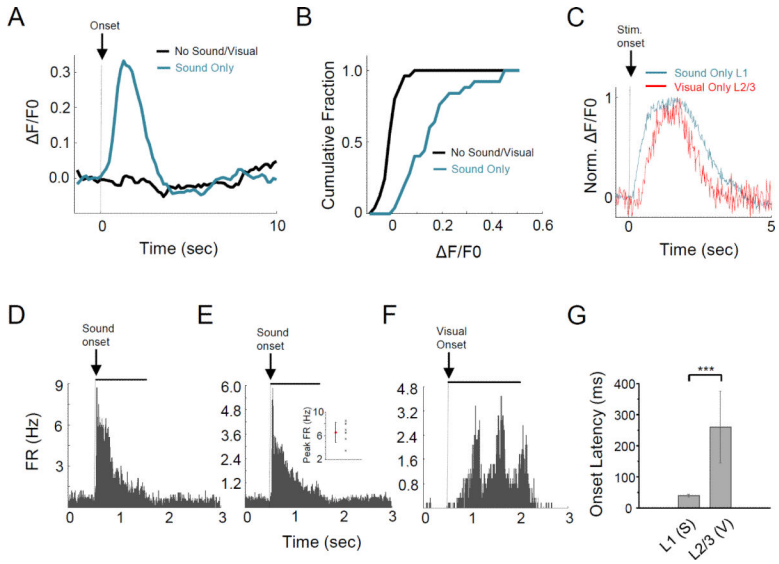

Figure 6. Sound alone excites L1 neurons

(A) Average Ca^{2+} response trace of an example L1 neuron to sound stimulation alone (blue). Black, trigger only with zero sound output.

(B) Cumulative distribution of L1 neurons' peak Ca^{2+} response amplitudes to sound (n = 27 cells).

(C) Population Ca^{2+} response of 5 L1 neurons to sound stimulation alone (blue) and that of 5 L2/3 neurons to visual stimulation alone (red) imaged in the same mouse, both normalized to their respective peak amplitude.

(D) PSTH of spike responses of an example L1 neuron to sound stimulation alone. Solid black line marks the sound duration.

(E) Average PSTH for 6 L1 neurons. Inset, peak firing rates of individual neurons and their mean. Bar = SD.

(F) PSTH for responses of an example L2/3 neuron to visual stimulation alone. Black bar marks the duration of moving gratings.

(G) Average onset latencies of L1 neuron spike responses to sound stimulation and L2/3 neuron spike responses to gratings. Onset latency was determined from the PSTH of an individual cell by the time when spike rate exceeds the average baseline firing rate plus two standard deviations of baseline fluctuations. Bar = SD. **p < 0.01, t-test, n = 6 for L1 and 14 for L2/3.

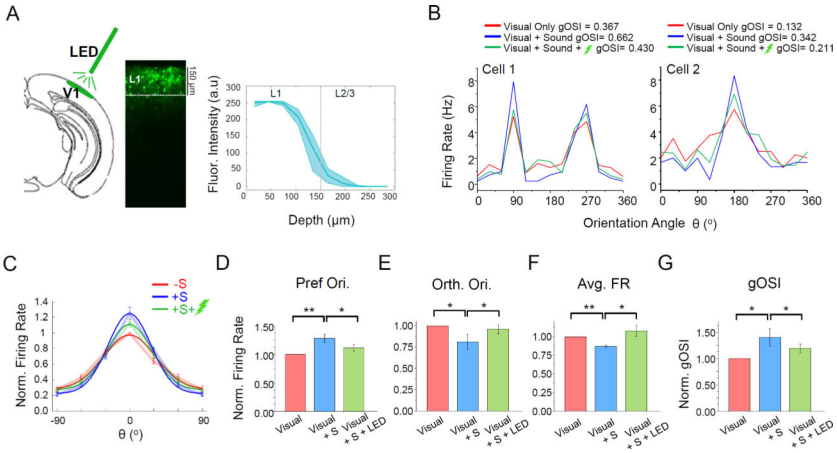

Figure 7. Suppressing L1 neurons reduces the sound-induced OS sharpening effect

(A) Left, schematic diagram of the experimental setup. Green LED illumination was applied to V1 surface. Middle, expression pattern of ArchT-GFP in a GAD2-Cre mouse. Right, average fluorescence intensity of ArchT-GFP at different depths across animals examined (n = 4). Shade = 95% confidence level.

(B) Tuning curves of evoked firing rate for two example L2/3 pyramidal neurons. Red, blue and green lines represent visual stimulation only, visual plus sound stimulation, visual plus sound plus inhibiting L1, respectively.

(C) Average tuning curves in three different conditions across all the cells recorded (n = 11).

(D-G) Normalized evoked firing rates at the preferred (D) and orthogonal orientation (E), average evoked firing rates across all orientations (F), as well gOSI values (G) under the three different conditions. Data represent mean ± SEM. *, $p < 0.05$; **, $p < 0.01$, one-way ANOVA with repeated measures.

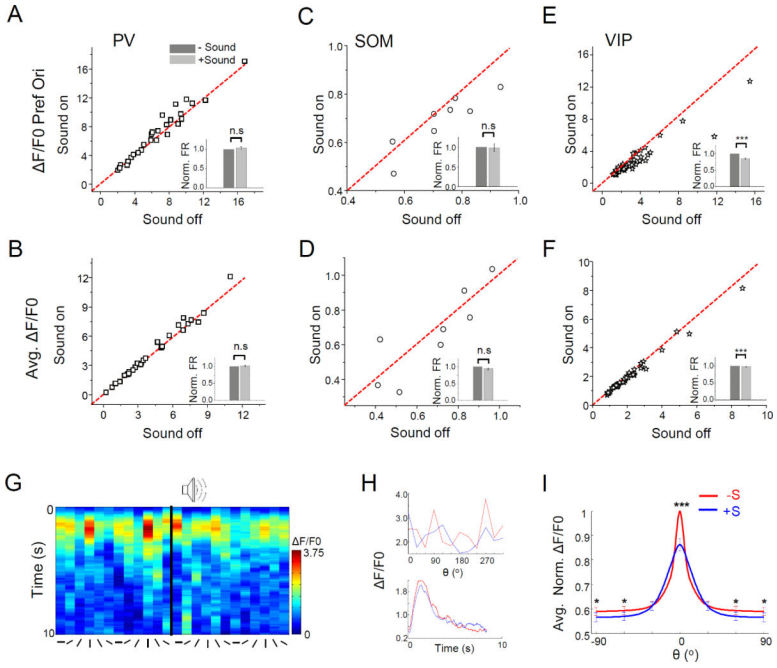

Figure 8. Effects of sound stimulation on PV, SOM and VIP neurons in L2/3

(A-B) Plots of peak Ca^{2+} response amplitude at the preferred orientation (A) and of average response amplitude across orientations (B) in the sound-on versus sound-off conditions for PVneurons (n = 30). Insets, mean normalized response amplitudes. Bar = SEM. “n.s.”, $p > 0.05$, Wilcoxon signed-rank test.

(C-D) Plots of response amplitudes for SOM neurons (n = 8). Data are presented in the same manner as in (A-B). “n.s.”, $p > 0.05$, paired t-test.

(E-F) Plots of response amplitudes for VIP neurons (n = 40). ***, $p < 0.001$, Wilcoxon signed-rank test (E) or paired t-test (F).

(G) Color map of ΔF/F0 for an example L2/3 VIP neuron.

(H) Top, tuning curves of peak ΔF/F0. Bottom, time-dependent ΔF/F0 averaged across all orientations for the same cell shown in (G). Red, visual stimulation alone; blue, visual plus sound stimulation.

(I) Average of Gaussian fits of tuning curves of individual VIP neurons. * $p < 0.05$; *** $p < 0.001$, Wilcoxon signed-rank test.