A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex

Andreas J Keller[1,3], Morgane M Roth[1,3], Matthew S. Caudil[2,3], Mario Dipoppa[4], Kenneth D. Miller[4,5], and Massimo Scanziani[1,2,3]

[1] Department of Physiology, University of California San Francisco, San Francisco, California, USA.

[2] Center for Neural Circuits and Behavior, Neurobiology Section and Department of Neuroscience, University of California San Diego, La Jolla, California, USA.

[3] Howard Hughes Medical Institute, University of California San Francisco, San Francisco, California, USA.

[4] Center for Theoretical Neuroscience, College of Physicians and Surgeons and Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY

[5] Dept. of Neuroscience, Swartz Program in Theoretical Neuroscience, Kavli Institute for Brain Science, College of Physicians and Surgeons and Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY

SUMMARY

Context strongly guides perception by influencing the saliency of sensory stimuli. Accordingly, sensory neurons respond differently to the same stimulus depending on the context in which the stimulus is embedded. In the visual system, the relationship between the features of the stimulus and those of the surround determine how excitatory neurons are modulated. Their responses are suppressed when center and surround share the same features but not when they are different. The mechanisms that relieve surround suppression and thereby generate contextual modulation remain unclear. Here we show that the disinhibitory circuit consisting of somatostatin-expressing (SOM) and vasoactive-intestinal-peptide-expressing (VIP) inhibitory neurons contextually modulates excitatory neurons in mouse primary visual cortex. When the stimulus and the surround share the same features, VIP neurons are inactive and SOM neurons can suppress excitatory neurons. However, when the features of the stimulus differ from those of the surround, VIP neurons are activated, thereby inhibiting SOM neurons and relieving excitatory neurons from suppression. Optogenetic silencing of VIP neurons, via disinhibition of SOM neurons, reduces the contextual modulation of excitatory neurons. We have identified a canonical cortical disinhibitory circuit which contributes to contextual modulation of excitatory neurons, and may thereby regulate perceptual saliency.

INTRODUCTION

The perception of a sensory stimulus is markedly influenced by the context in which the stimulus is embedded. In the visual system, the context is the visual scene surrounding the stimulus. Through the

influence of its surround, the same visual stimulus may be perceived as more or less salient, allowing it to pop out or merge with the rest of the visual scene (Bergen and Julesz, 1983; Lamme, 1995; Treisman and Gelade, 1980). This aspect of sensory processing represents a fundamental computation to extract meaning from visual scenes.

Consistent with perceptual phenomena, the stimulus evoked responses of neurons are modulated by the surround. This contextual modulation occurs at several stages of the visual system including the retina (Alitto and Usrey, 2008; Chiao and Masland, 2003; Huang et al., 2019; McIlwain, 1964; Ölveczky et al., 2003; Solomon, 2006), the thalamus (Alitto and Usrey, 2008; Jones et al., 2012, 2015; Levick et al., 1972), and the visual cortex (Alexander and Van Leeuwen, 2010; Angelucci et al., 2017; Fitzpatrick, 2000; Kapadia et al., 2000; Knierim and van Essen, 1992; Rossi et al., 2001; Schnabel et al., 2018; Slllito et al., 1995) progressively increasing the complexity of the contextualized spatial features.

Excitatory neurons in primary visual cortex (V1) are activated by patches of gratings centered on their classical feedforward receptive field (ffRF; Hubel and Wiesel, 1962). A well-established example of contextual modulation is surround suppression, which occurs when the stimulus features in the neuron's ffRF are also present away from its ffRF, i.e. in its surround (Blakemore and Tobin, 1972; Hubel and Wiesel, 1965; Kapadia et al., 1999; Knierim and van Essen, 1992; Nelson and Frost, 1978). It has been shown that the mechanisms for surround suppression involve feedback connections (Angelucci et al., 2017; Nurminen et al., 2018; Vangeneugden et al., 2019), interlaminar connections (Bolz and Gilbert, 1986) and specific subtypes of inhibitory neurons (Adesnik et al., 2012; Haider et al., 2010). The tuning properties of somatostatin-expressing (SOM) inhibitory neurons (Adesnik et al., 2012) and the fact that they connect to nearly all surrounding excitatory neurons (Fino et al., 2013) make them ideal to contribute to surround suppression. Indeed, by functional elimination of SOM neurons, excitatory neurons are relieved from the suppression mediated by large stimuli covering center and surround (Adesnik et al., 2012).

In a different context, obtained by changing the orientation of the gratings in the surround, surround suppression in excitatory neurons is weakened and their responses may even be facilitated (Self et al., 2014; Slllito et al., 1995; Walker et al., 1999). Yet, the mechanisms that eliminate surround suppression remain elusive. Surround suppression and facilitation are fundamental computations that modulate the responses of sensory neurons based on the context.

To investigate the mechanisms of contextual modulation, we presented visual stimuli with different visual contexts to awake mice while imaging calcium responses in layer 2/3 excitatory neurons of V1. To verify

that the contextual modulation of excitatory neurons is not simply inherited from earlier stages of visual processing, we performed similar experiments in layer 4, the main feedforward input layer of cortex. We determined the sources of contextual modulation by recording from the three main inhibitory neuron subtypes, vasoactive-intestinal-peptide-expressing (VIP), SOM, and parvalbumin-expressing (PV) neurons. We show that while VIP and PV neurons are modulated by context in a similar way to excitatory neurons, SOM neurons are modulated conversely. Furthermore, we show that when functionally eliminating VIP neurons, the main inhibitory source of input to SOM neurons, SOM neurons lose their context specific modulation and in turn excitatory neurons reduce their contextual modulation. These results are consistent with a cortical disinhibitory circuit responsible for contextual modulation in excitatory neurons.

RESULTS

Excitatory neurons are increasingly modulated by context along the laminar processing hierarchy

To assess the contextual modulation in sensory cortex, we recorded from layer 2/3 excitatory neurons in primary visual cortex (V1) in awake head-fixed mice with two-photon calcium imaging (Fig. 1A). To estimate the contextual modulation of the neurons, we compared their responses to a small grating patch (20° in diameter) presented alone ("center stimulus"), with an iso-oriented surround ("iso stimulus"), or with a cross-oriented surround ("cross stimulus"). The orientation and location of the center stimulus was matched with the ffRF properties of the neurons (see Methods). Consistent with previous studies (Self et al., 2014; Slllito et al., 1995; Walker et al., 1999), neurons were strongly suppressed by an iso but not by a cross stimulus. This contextual modulation persisted when the orientation of the center stimulus did not match the neuron's orientation preference (data not shown). Therefore, we included neurons independently of their orientation preference hereafter (Fig. 1). Some neurons not only recovered their activity to the same level of responses observed to the center stimulus but responded stronger than for the center stimulus alone and therefore much stronger than for an iso stimulus (Fig. 1B, C). To compare the modulation by an iso surround to that of a cross surround, we defined a surround modulation index for each neuron (Fig. 1D; see inset for definition). The distribution of modulation indexes of excitatory neurons was skewed to positive values indicating that their responses were stronger to the cross than to the iso stimulus. Overall, excitatory neurons in layer 2/3 were strongly modulated by context, i.e. the strength of their responses depended on the features of the surround.

We then asked to which extent the contextual modulation of excitatory layer 2/3 neurons was inherited from earlier stages of cortical processing. To this end, we imaged the activity of excitatory neurons in layer

4, the main cortical feedforward input layer, by conditionally expressing GCaMP6f (Fig. 1A bottom), and presented center, iso and cross stimuli at the ffRF location of the imaged neurons. While layer 2/3 neurons, on average, were only suppressed by the iso stimulus, layer 4 neurons showed suppression in both contexts (Fig. 1B, bottom; Fig. 1C). Indeed, the surround modulation indices of layer 4 neurons were significantly smaller than those of layer 2/3 neurons (Fig. 1D). Thus, contextual modulation of layer 2/3 neurons is unlikely to be entirely inherited from layer 4, and may rely on local circuitry.

We hypothesized that contextual modulation in layer 2/3 neurons is computed within layer 2/3. Inhibitory neurons in visual cortex have tuning properties that often differ from those of excitatory neurons (Adesnik et al., 2012; Atallah et al., 2012; Keller and Martin, 2015; Kerlin et al., 2010; Khan et al., 2018; Millman et al., 2019; Niell and Stryker, 2008; Pakan et al., 2016; Sohya et al., 2007). Could context modulate excitatory neurons through inhibition and disinhibition of specific neuronal subtypes?

Complementary contextual modulation in SOM and VIP neurons

To investigate whether contextual modulation is neuron-type specific, we characterized the responses to center, iso, and cross stimulus for the three major subtypes of inhibitory neurons, vasoactive-intestinal-peptide- (VIP), parvalbumin- (PV), and somatostatin-expressing (SOM) neurons. To this end, we conditionally expressed GCaMP6f in each of these three classes and repeated the protocol used for excitatory neurons. VIP and PV neurons were facilitated by a cross and suppressed by an iso stimulus (VIP, Fig. 2A-C; PV, data not shown), similarly to what we observed in excitatory neurons. Strikingly, contextual modulation of SOM neurons was opposite to that of VIP and PV neurons, and importantly, opposite to the contextual modulation of excitatory neurons (Fig. 2D-F).

These results reveal a neuron-type specific contextual modulation. This raises important questions on the inhibitory circuitry responsible for the contextual modulation in excitatory neurons. Indeed, the direct inhibition from VIP and PV neurons onto excitatory neurons is unlikely to be responsible for their contextual modulation. However, the reduction of SOM neuron responses to the cross compared to the center stimulus could lead to the facilitation observed in excitatory neurons. Moreover, it has been shown that functional elimination of SOM neurons relieves excitatory neurons from the suppression observed in response to the iso stimulus (Adesnik et al., 2012). Together, SOM neurons could directly cause the contextual modulation of excitatory neurons for both iso and cross stimuli.

Which mechanism leads to the suppression of SOM responses to cross stimuli? The main inhibitory input onto SOM neurons is provided by VIP neurons (Pfeffer et al., 2013). Moreover, VIP neurons are strongly

activated by cross and suppressed by an iso stimulus (Fig. 2A-C). Hence, both connectivity and physiology of VIP make them ideal to contextually modulate SOM neurons. We therefore propose a model in which, in response to a center stimulus, the network is in a balanced state (Fig. 2G, left) but in response to an iso stimulus, the circuit is dominated by SOM neurons, which inhibit all other neuron types (Fig. 2G, middle). For the cross stimulus, however, VIP neurons are strongly activated and inhibit SOM neurons which leads to disinhibition of excitatory neurons (Fig. 2G, right). A central prediction of this model is that the functional elimination of VIP neurons would, via release of inhibition in SOM neurons, reduce the contextual modulation in excitatory neurons, leading to suppression in both contexts.

Reciprocal inhibition between VIP and SOM contributes to contextual modulation in excitatory neurons

Does the functional elimination of VIP neurons influence the contextual modulation in excitatory neurons? We expressed an inhibitory opsin conditionally in VIP neurons and GCaMP7f unconditionally in order to manipulate VIP activity while recording the activity in identified VIP neurons and in putative excitatory neurons (Fig 3A). By combining calcium imaging and optogenetic silencing, we recorded responses of VIP neurons to center, iso and cross stimuli with and without silencing (see Methods). With this approach, we were able to significantly reduce the stimulus evoked responses in VIP neurons (data not shown). We specifically tested if by silencing VIP neurons excitatory neurons would reduce their responses to cross more than to iso stimuli and therefore abolish contextual modulation in excitatory neurons. Since excitatory neurons were already strongly suppressed in the iso condition, we reduced the contrast of all stimuli to 50%, to evoke less iso-surround suppression (Kapadia et al., 1999). When silencing VIP neurons, the responses of putative excitatory neurons to cross stimuli were significantly more suppressed than those to iso stimuli (Fig. 3B, C; also true for 100% stimuli, data not shown). During this manipulation, excitatory neurons shifted their surround modulation index towards zero (Fig. 3D), indicating that they were increasingly independent of the specific context. In addition, silencing VIP neurons had a strong suppressive effect on the baseline activity of excitatory neurons, suggesting further a disinhibitory role of VIP neurons (data not shown). SOM neurons are one of the main targets of VIP neurons and are likely to be the missing link between VIP neurons and excitatory neurons.

To test if the disinhibition of excitatory neurons involves SOM neurons, we silenced VIP neurons while recording in SOM neurons expressing GCaMP6s (using a cre-lox system combined with a FLP-FRT system; Fig. 3E). When silencing VIP neurons, the responses of SOM neurons to the cross stimulus were strongly increased while the responses to the iso stimulus was less affected (Fig. 3F, G). Accordingly, the surround modulation index of SOM neurons significantly increased upon VIP neuron silencing and its distribution

was no longer significantly different from a distribution around zero (Fig. 3H). In conclusion, when VIP neurons were functionally eliminated, SOM neurons responded equally in both contexts indicating that they were no longer modulated by context.

Taken together, these results demonstrate that the VIP-SOM disinhibitory circuit contributes to contextual modulation in excitatory neurons.

DISCUSSION

This study provides a mechanism for contextual modulation in the primary visual cortex and reveals a disinhibitory circuit as a key mediator. Using imaging and causal manipulations, we find that the relationship between SOM and VIP inhibitory neurons determine the response profiles of layer 2/3 excitatory neurons in V1. When a uniform full-field stimulus is presented, VIP neurons are silent, SOM neurons dominate the network and inhibit excitatory neurons (see also Adesnik et al., 2012). With a rupture in orientation between center and surround, VIP neurons are excited, inhibiting SOM neurons and effectively relieving excitatory neurons from SOM inhibition.

Sources of inputs onto inhibitory neurons

The connectivity motifs between inhibitory neurons has been previously described (Jiang et al., 2015; Pfeffer et al., 2013; Pi et al., 2013) and supports our findings. SOM neurons connect to all other subtypes while VIP neurons preferentially inhibit SOM neurons. In addition, SOM neurons participate in the surround suppression observed in excitatory neurons by integrating over a large cortical area (Adesnik et al., 2012). Our study demonstrates that the interaction between SOM and VIP neurons contributes to the contextual modulation observed in excitatory neurons. When SOM neurons prevail over VIP neurons, excitatory neurons are inhibited, i.e. surround suppressed. Conversely, when VIP neurons prevail over SOM neurons, excitatory neurons are relieved from suppression, i.e. cross-surround facilitated. The mechanisms that tip the balance in favor of VIP neurons remain to be found. VIP neurons could receive additional inputs, such as modulatory inputs (Kawaguchi, 1997; Paspalas and Papadopoulos, 2001; van Versendaal and Levelt, 2016) as well as long range cortical and thalamic inputs (van Versendaal and Levelt, 2016; Wall et al., 2016; Zhang et al., 2014). In addition, they have been suggested to be the target of feedback connections (van Versendaal and Levelt, 2016) and, in accordance, we previously found that VIP neurons have a second receptive field driven by feedback (Keller et al., 2020). All these sources could drive VIP neurons in response to cross stimuli.

Feedback drive

In a recent study, we found that VIP neurons strongly respond to the context alone, when their feedforward receptive field is not stimulated (Keller et al., 2020). Interestingly, while excitatory and PV neurons also have a feedback receptive field, SOM neurons do not. Moreover, orientation preferences of the feedforward and feedback receptive fields are independent. The feedback receptive field could therefore be an interesting source for contextual modulation. We hypothesize that the orientation preferences of the feedforward and the feedback receptive fields of VIP neurons are not aligned, so that the simultaneous stimulation of both receptive fields would provide the strong drive supporting cross-orientation facilitation. However, we cannot exclude that contextual modulation could partially be inherited from earlier stages of the visual system. Indeed, we find that layer 4 excitatory neurons are modulated by context, albeit weakly (Fig. 1). Moreover, Self and colleagues show that the contextual modulation in layer 4 is resistant to layer 2/3 silencing (Self et al., 2014).

Conclusions

Contextual modulation represents a fundamental computation to extract meaning from visual scenes. It could support many perceptual phenomena, such as pop-out effects, figure-ground segregation, detection of borders, and object detection (Angelucci et al., 2017; Bergen and Julesz, 1983; Jones et al., 2001; Kapadia et al., 2000; Knierim and van Essen, 1992; Lamme, 1995; Rossi et al., 2001; Schnabel et al., 2018; Seriès et al., 2003; Treisman and Gelade, 1980). Furthermore, the dichotomy between surround suppression and cross-orientation facilitation is consistent with a predictive processing framework (Bastos et al., 2012; Keller and Mrsic-Flogel, 2018; Rao and Ballard, 1999). Based on natural statistics of the visual environment, the spatial features in a small patch of visual world are more likely to be similar to the spatial features in the adjacent patches. If the spatial prediction of the center based on the surround is correct the response of that neuron can be suppressed, i.e. surround suppression, as there is no need to transmit a signal that is accurately predicted. On the other hand, if a rupture or a dissimilarity is present between the center and the surround, the prediction of the center based on the surround is incorrect and the signal of the neuron will not be suppressed but passed along, i.e. cross-orientation facilitation. In conclusion, predictive processing is a strong framework for contextual modulation of visual responses in cortical circuits.

REFERENCES

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

Alexander, D.M., and Van Leeuwen, C. (2010). Mapping of contextual modulation in the population response of primary visual cortex. Cogn. Neurodyn. *4*, 1–24.

Alitto, H.J., and Usrey, W.M. (2008). Origin and Dynamics of Extraclassical Suppression in the Lateral Geniculate Nucleus of the Macaque Monkey. Neuron *57*, 135–146.

Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., and Bressloff, P.C. (2017). Circuits and Mechanisms for Surround Modulation in Visual Cortex. Annu. Rev. Neurosci. *40*, 425–451.

Atallah, B.V., Bruns, W., Carandini, M., and Scanziani, M. (2012). Parvalbumin-Expressing Interneurons Linearly Transform Cortical Responses to Visual Stimuli. Neuron *73*, 159–170.

Attinger, A., Wang, B., and Keller, G.B. (2017). Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. Cell *169*, 1291-1302.e14.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical Microcircuits for Predictive Coding. Neuron *76*, 695–711.

Bergen, J.R., and Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. Nature *303*, 696–698.

Blakemore, C., and Tobin, ElisabethA. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex. Exp. Brain Res. *15*.

Bolz, J., and Gilbert, C.D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. Nature *320*, 362–365.

Chiao, C.-C., and Masland, R.H. (2003). Contextual tuning of direction-selective retinal ganglion cells. Nat. Neurosci. *6*, 1251–1252.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron *56*, 43–57.

Fino, E., Packer, A.M., and Yuste, R. (2013). The Logic of Inhibitory Connectivity in the Neocortex. The Neuroscientist *19*, 228–237.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. Curr. Opin. Neurobiol. *10*, 438–443.

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J.A., and McCormick, D.A. (2010). Synaptic and Network Mechanisms of Sparse and Reliable Visual Cortical Activity during Nonclassical Receptive Field Stimulation. Neuron *65*, 107–121.

Huang, X., Rangel, M., Briggman, K.L., and Wei, W. (2019). Neural mechanisms of contextual modulation in the retinal direction selective circuit. Nat. Commun. *10*, 2431.

Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. *160*, 106–154.

Hubel, D.H., and Wiesel, T.N. (1965). RECEPTIVE FIELDS AND FUNCTIONAL ARCHITECTURE IN TWO NONSTRIATE VISUAL AREAS (18 AND 19) OF THE CAT. J. Neurophysiol. *28*, 229–289.

Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. Science *350*, aac9462–aac9462.

Jones, H.E., Grieve, K.L., Wang, W., and Sillito, A.M. (2001). Surround Suppression in Primate V1. J. Neurophysiol. *86*, 2011–2028.

Jones, H.E., Andolina, I.M., Ahmed, B., Shipp, S.D., Clements, J.T.C., Grieve, K.L., Cudeiro, J., Salt, T.E., and Sillito, A.M. (2012). Differential Feedback Modulation of Center and Surround Mechanisms in Parvocellular Cells in the Visual Thalamus. J. Neurosci. *32*, 15946–15951.

Jones, H.E., Andolina, I.M., Shipp, S.D., Adams, D.L., Cudeiro, J., Salt, T.E., and Sillito, A.M. (2015). Figure-ground modulation in awake primate thalamus. Proc. Natl. Acad. Sci. *112*, 7085–7090.

Kapadia, M.K., Westheimer, G., and Gilbert, C.D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. Proc. Natl. Acad. Sci. *96*, 12073–12078.

Kapadia, M.K., Westheimer, G., and Gilbert, C.D. (2000). Spatial Distribution of Contextual Interactions in Primary Visual Cortex and in Visual Perception. J. Neurophysiol. *84*, 2048–2062.

Kawaguchi, Y. (1997). Selective Cholinergic Modulation of Cortical GABAergic Cell Subtypes. J. Neurophysiol. *78*, 1743–1747.

Keller, A.J., and Martin, K.A.C. (2015). Local Circuits for Contrast Normalization and Adaptation Investigated with Two-Photon Imaging in Cat Primary Visual Cortex. J. Neurosci. *35*, 10078–10087.

Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron *100*, 424–435.

Keller, A.J., Roth, M.M., and Scanziani, M. (2020). Neurons in Visual Cortex are Driven by Feedback Projections when their Feedforward Sensory Input is Missing (bioRxiv).

Kerlin, A.M., Andermann, M.L., Berezovskii, V.K., and Reid, R.C. (2010). Broadly Tuned Response Properties of Diverse Inhibitory Neuron Subtypes in Mouse Visual Cortex. Neuron *67*, 858–871.

Khan, A.G., Poort, J., Chadwick, A., Blot, A., Sahani, M., Mrsic-Flogel, T.D., and Hofer, S.B. (2018). Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. Nat. Neurosci. *21*, 851–859.

Knierim, J.J., and van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol. *67*, 961–980.

Lamme, V. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. J. Neurosci. *15*, 1605–1615.

Leinweber, M., Zmarz, P., Buchmann, P., Argast, P., Hübener, M., Bonhoeffer, T., and Keller, G.B. (2014). Two-photon Calcium Imaging in Mice Navigating a Virtual Reality Environment. J. Vis. Exp. 50885.

Levick, W.R., Cleland, B.G., and Dubin, M.W. (1972). Lateral geniculate neurons of cat: Retinal inputs and physiology. Invest. Ophthalmol. *11*, 10.

McIlwain, J.T. (1964). RECEPTIVE FIELDS OF OPTIC TRACT AXONS AND LATERAL GENICULATE CELLS: PERIPHERAL EXTENT AND BARBITURATE SENSITIVITY. J. Neurophysiol. *27*, 1154–1173.

Millman, D.J., Ocker, G.K., Caldejon, S., Kato, I., Larkin, J.D., Lee, E.K., Luviano, J., Nayan, C., Nguyen, T.V., North, K., et al. (2019). VIP interneurons selectively enhance weak but behaviorally-relevant stimuli (bioRxiv).

Nelson, J.I., and Frost, B.J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. Brain Res. *139*, 359–365.

Niell, C.M., and Stryker, M.P. (2008). Highly Selective Receptive Fields in Mouse Visual Cortex. J. Neurosci. *28*, 7520–7536.

Nurminen, L., Merlin, S., Bijanzadeh, M., Federer, F., and Angelucci, A. (2018). Top-down feedback controls spatial summation and response amplitude in primate visual cortex. Nat. Commun. *9*, 2281.

Ölveczky, B.P., Baccus, S.A., and Meister, M. (2003). Segregation of object and background motion in the retina. Nature *423*, 401–408.

Pakan, J.M., Lowe, S.C., Dylda, E., Keemink, S.W., Currie, S.P., Coutts, C.A., and Rochefort, N.L. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. ELife *5*, e14985.

Paspalas, C.D., and Papadopoulos, G.C. (2001). Serotoninergic afferents preferentially innervate distinct subclasses of peptidergic interneurons in the rat visual cortex. Brain Res. *891*, 158–167.

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat. Neurosci. *16*, 1068–1076.

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. Nature *503*, 521–524.

Rao, R.P.N., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87.

Rossi, A.F., Desimone, R., and Ungerleider, L.G. (2001). Contextual Modulation in Primary Visual Cortex of Macaques. J. Neurosci. *21*, 1698–1709.

Schnabel, U.H., Bossens, C., Lorteije, J.A.M., Self, M.W., Op de Beeck, H., and Roelfsema, P.R. (2018). Figure-ground perception in the awake mouse and neuronal activity elicited by figure-ground stimuli in primary visual cortex. Sci. Rep. *8*, 17800.

Self, M.W., Lorteije, J.A.M., Vangeneugden, J., van Beest, E.H., Grigore, M.E., Levelt, C.N., Heimel, J.A., and Roelfsema, P.R. (2014). Orientation-Tuned Surround Suppression in Mouse Visual Cortex. J. Neurosci. *34*, 9290–9304.

Seriès, P., Lorenceau, J., and Frégnac, Y. (2003). The “silent” surround of V1 receptive fields: theory and experiments. J. Physiol.-Paris *97*, 453–474.

Slllito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J., and Davls, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. Nature *378*, 492–496.

Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K., and Tsumoto, T. (2007). GABAergic Neurons Are Less Selective to Stimulus Orientation than Excitatory Neurons in Layer II/III of Visual Cortex, as Revealed by In Vivo Functional Ca2+ Imaging in Transgenic Mice. J. Neurosci. *27*, 2145–2149.

Solomon, S.G. (2006). Suppressive Surrounds and Contrast Gain in Magnocellular-Pathway Retinal Ganglion Cells of Macaque. J. Neurosci. *26*, 8715–8726.

Treisman, A.M., and Gelade, G. (1980). A feature-integration theory of attention. Cognit. Psychol. *12*, 97–136.

Vangeneugden, J., van Beest, E.H., Cohen, M.X., Lorteije, J.A.M., Mukherjee, S., Kirchberger, L., Montijn, J.S., Thamizharasu, P., Camillo, D., Levelt, C.N., et al. (2019). Activity in Lateral Visual Areas Contributes to Surround Suppression in Awake Mouse V1. Curr. Biol. *29*, 4268-4275.e7.

van Versendaal, D., and Levelt, C.N. (2016). Inhibitory interneurons in visual cortical plasticity. Cell. Mol. Life Sci. *73*, 3677–3691.

Walker, G.A., Ohzawa, I., and Freeman, R.D. (1999). Asymmetric Suppression Outside the Classical Receptive Field of the Visual Cortex. J. Neurosci. *19*, 10536–10553.

Wall, N.R., De La Parra, M., Sorokin, J.M., Taniguchi, H., Huang, Z.J., and Callaway, E.M. (2016). Brain-Wide Maps of Synaptic Input to Cortical Interneurons. J. Neurosci. *36*, 4000–4009.

Xue, M., Atallah, B.V., and Scanziani, M. (2014). Equalizing excitation–inhibition ratios across visual cortical neurons. Nature *511*, 596–600.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. Science *345*, 660–665.

Acknowledgements. This project was supported by the NIH grant U19NS107613, the Howard Hughes Medical Institute and the Swiss National Science Foundation grants P300PA_177882 and P2EZP3_162284 to A.J.K and P300PA_177898 to M.M.R, the Gatsby Charitable Foundation and NSF NeuroNex Award DBI-1707398 to K.D.M. and M.D.

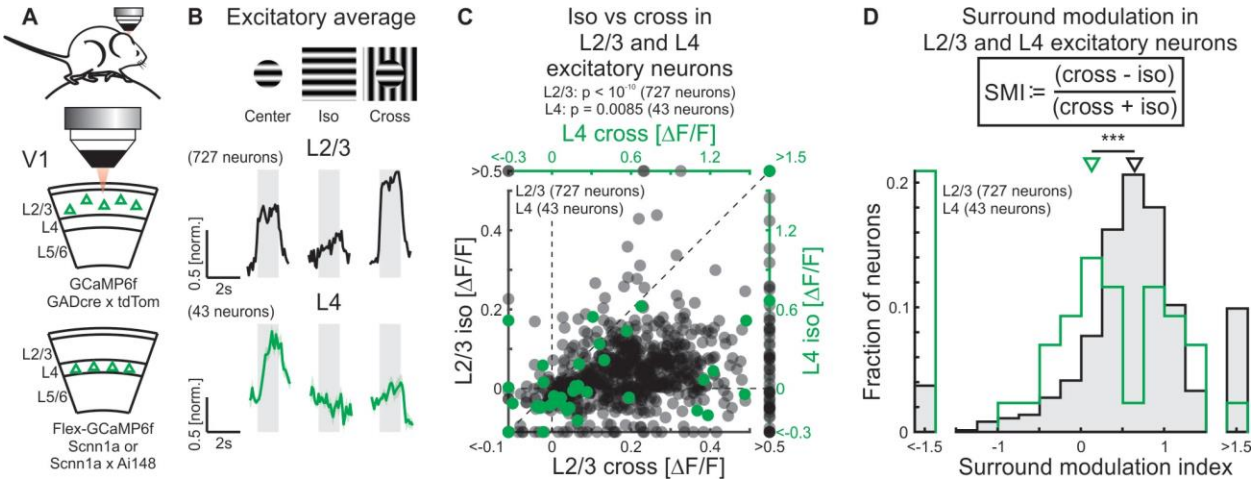

Figure 1. Contextual modulation in excitatory neurons

(A) Visual stimuli are presented to awake mice while imaging calcium responses in layer 2/3 or layer 4 excitatory neurons of primary visual cortex (V1).

(B) Top: Schematic of small grating patch (20° in diameter) presented alone (center), with an iso-oriented surround (iso), or with a cross-oriented surround (cross). Bottom: Population-averaged calcium responses of layer 2/3 and layer 4 excitatory neurons to center, iso, and cross stimuli (727 neurons in 9 mice and 43 neurons in 5 mice in layer 2/3 and layer 4, respectively). Responses of every neuron were normalized by their maximum response among the three stimuli. Here and in all other figures shaded areas are periods of stimulus presentation. Note the strong response to cross compared to iso in L2/3 and the suppression to both iso and cross in L4.

(C) Scatter plot of the responses to iso and cross. Wilcoxon sign-rank test; layer 2/3: $p < 10^{-10}$; layer 4: $p = 0.0085$; same neurons as in (B).

(D) Surround modulation index (SMI) was computed as the difference divided by the sum of cross and iso stimuli (see box). Here and in all figures triangles above histograms indicate median. Wilcoxon rank-sum test; $p = 3.6 \times 10^{-4}$; same neurons as in (B).

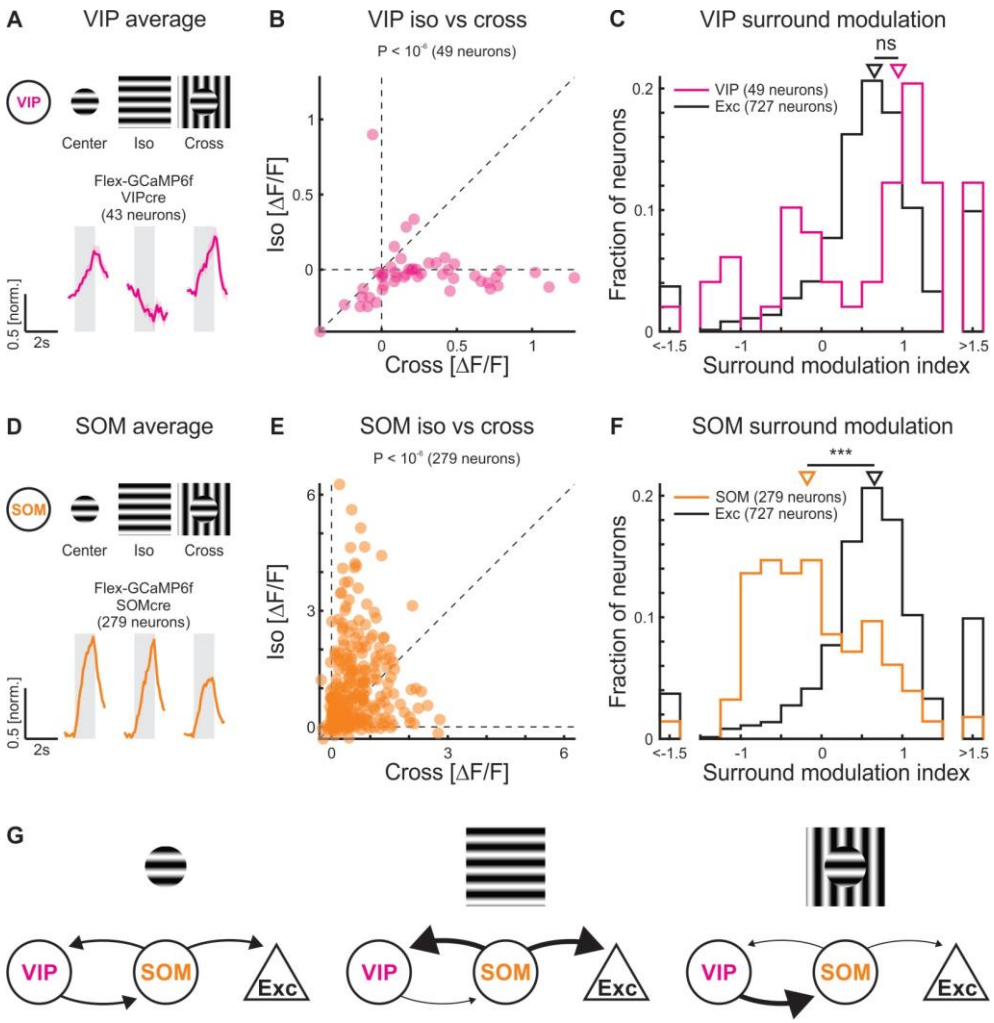

Figure 2. Contextual modulation in inhibitory neurons

(A) Population-averaged calcium responses of vasoactive-intestinal-peptide-expressing (VIP) inhibitory neurons to center, iso, and cross stimuli (49 neurons in 6 mice). Responses of every neuron were normalized by their maximum response among these three stimuli.

(B) Scatter plot of the responses to iso and cross stimuli. Wilcoxon sign-rank test; $p < 10^{-6}$; 49 neurons in 6 mice.

(C) SMI distributions of VIP compared to layer 2/3 excitatory neurons. Wilcoxon rank-sum test; $p = 0.49$; 49 VIP and 727 layer 2/3 excitatory neurons.

(D-F) As above, but for somatostatin-expressing (SOM) inhibitory neurons (279 neurons in 13 mice).

(E) Note that SOM neurons are significantly more activated by iso than by cross stimuli, opposite to other neuron types. Wilcoxon sign-rank test; $p < 10^{-6}$; 279 neurons in 13 mice.

(F) While the SMI distribution of excitatory neurons is skewed to the right that of SOM neurons is skewed to the left. Wilcoxon rank-sum test; $p < 10^{-10}$; 279 SOM and 727 layer 2/3 excitatory neurons.

(G) Schematic model of contextual modulation of excitatory neurons through the cooperation of VIP and SOM neurons. Left: The circuit is in a balanced state in response to the center stimulus. Middle: In response to an iso stimulus, SOM neurons are strongly driven and inhibit both VIP and excitatory neurons. Right: In response to the cross stimulus, VIP neurons are strongly driven and inhibit SOM neurons. Suppression of SOM neurons in turn disinhibits excitatory neurons.

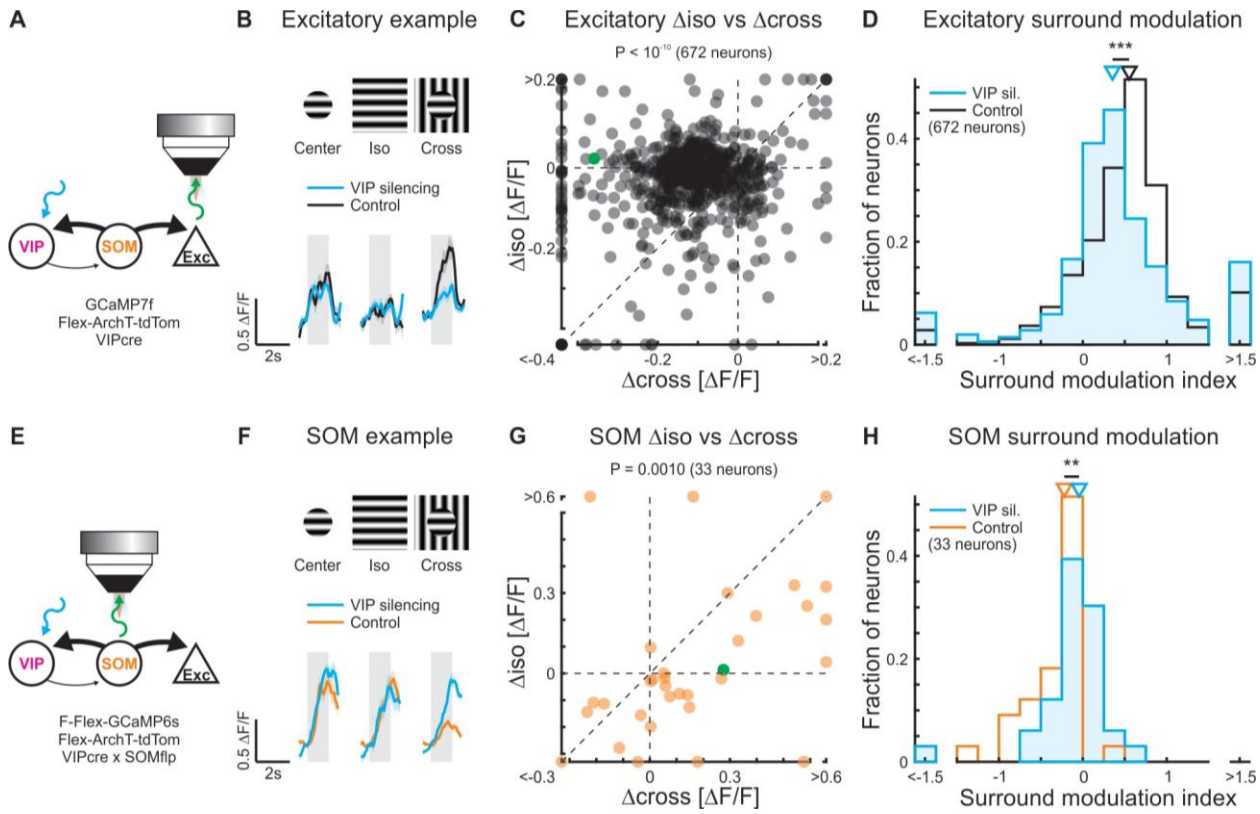

Figure 3. VIP and SOM neurons cooperatively contribute to contextual modulation in excitatory neurons

(A) Experimental setup. We conditionally expressed an inhibitory opsin in VIP neurons and a calcium indicator was expressed unconditionally.

(B) Trial-averaged calcium responses of a putative excitatory neuron with and without silencing of VIP neurons. Here, stimuli were presented at 50% contrast (similar responses to 100% stimuli, data not shown).

(C) Scatter plot of changes in putative excitatory responses due to VIP silencing. Note that by silencing VIP neurons, responses of excitatory neurons are significantly more reduced to cross compared to iso stimuli. Wilcoxon sign-rank test; $p < 10^{-10}$; 672 neurons in 6 mice. Green symbol represents the example neuron shown in (B).

(D) Upon silencing VIP neurons, the SMI distribution of putative excitatory neurons shifts significantly to the left, closer to zero. Wilcoxon sign-rank test; $p < 10^{-4}$; 672 neurons in 6 mice.

(E) Same as A, but calcium indicator conditionally expressed in SOM neurons.

(F-H) Same as (B-D), but for SOM neurons.

(G) Wilcoxon sign-rank test; $p = 0.0010$; 33 neurons in 6 mice.

(H) Wilcoxon sign-rank test; $p = 6.6 \times 10^{-4}$; 33 neurons in 6 mice.

METHODS

Animals. All experimental procedures were conducted in accordance with the regulation of the Institutional Animal Care and Use Committee of the University of California, San Francisco. Mice were of either sex and were of the following genotype:

Gad2-IRES-cre (GAD2$^{tm2(cre)Zjh}$; JAX:010802) × Ai14 (Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze}$; JAX:007914) for imaging of layer 2/3 excitatory neurons (5 mice; Figures 1, 2C);

Scnn1a-Tg3-cre (Tg(Scnn1a-cre)3Aibs/J; JAX:009613) and Scnn1a-Tg3-cre (Tg(Scnn1a-cre)3Aibs/J; JAX:009613) × Ai148 (Igs7$^{tm148.1(tetO-GCaMP6f,CAG-tTA2)Hze}$; JAX:030328) for imaging layer 4 excitatory neurons (5 mice and 1 mouse, respectively; Figure 1;

VIP-IRES-cre (Vip$^{tm1(cre)Zjh}$; JAX:010908) × Ai14 (Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze}$; JAX:007914) for imaging of layer 2/3 vasoactive-intestinal-peptide-expressing inhibitory neurons (VIP; 7 mice; Figures 2A-C);

VIP-IRES-cre (Vip$^{tm1(cre)Zjh}$; JAX:010908) for optogenetic manipulation of VIP neurons and imaging putative excitatory and VIP neurons (2 mice; Figure 3A-D);

VIP-IRES-cre (Vip$^{tm1(cre)Zjh}$; JAX:010908) × Sst-IRES-Flp (Sst$^{tm3.1(flpo)Zjh}$; JAX:028579) for optogenetic manipulation of VIP neurons and imaging either somatostatin-expressing (SOM) inhibitory neurons or putative excitatory and VIP neurons (9 mice; Figure 3);

Sst-IRES-cre (Sst$^{tm2.1(cre)Zjh}$; JAX:028864) × Ai14 (Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze}$; JAX:007914) for imaging of layer 2/3 SOM (13 mice; Figures 2D-F); and

PV-cre (Pvalb$^{tm1(cre)Arbr}$; JAX:017320) × Ai14 (Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze}$; JAX:007914) for imaging of layer 2/3 parvalbumin-expressing inhibitory neurons (PV; 9 mice; data not shown).

The mice were housed on a reverse light cycle (light/dark cycle: 12/12 hrs). At the start of the experiments, all mice were older than 2 months.

Viruses. We injected the following viruses:

AAV2/1.ef1a.GCaMP6f.WPRE (FMI Vector Core Facility),

AAV2/1.ef1a.DIO.GCaMP6f.WPRE (FMI Vector Core Facility),

AAV2/1.CAG.CGaMP6f (Janelia Vector Core),

AAV2/9.syn.GCaMP7f (Addgene),

AAV2/1.ef1a.fDIO.GCaMP6s (Janelia Vector Core),

AAV2/5.CBA.Flex.ArchT-tdTomato.WPRE.SV40 (University of Pennsylvania Vector Core),

AAV2/1.CAG.Flex.rc[Jaws-KGC-GFP-ER2] (Janelia Vector Core), and

AAV2/9.ef1a.F-Flex.tdTomato (Xue et al., 2014).

Viruses were diluted to use titers of approximately 5×10^{12} genome copies/ml and 50nl were injected at each injection site (3 to 5 sites per mouse) and each depth (2 from 350 to 200 mm below the pial surface).

Surgery. Mice were anesthetized with 2% isoflurane or with a mixture of Fentanyl (West-Ward Pharmaceuticals, 0.05 mg/kg), Midazolam (Akorn, 5.0 mg/kg) and Dexmedetomidine (Zoetis, 0.5 mg/kg), injected subcutaneously. Mice's body temperature was monitored and kept constant. To prevent the eyes from drying, a layer of lubricant ointment (Rugby) was applied. The skin above the skull was disinfected with povidone iodine. A craniotomy was made over the right visual cortex (3 to 4.5 mm in diameter) and viruses were injected with a micropump (UMP-3, World Precision Instruments) at a rate of 2 nl/s. The craniotomy was then sealed with a glass coverslip using cyanoacrylate glue and a headplate was attached. To reverse the anesthesia induced by the Fentanyl-Midazolam-Dexmedetomidine mixture, a mixture of Naloxone (Hospira, 1.2 mg/kg), Flumazenil (West-Ward Pharmaceuticals, 0.5 mg/kg), and Atipamezol (Zoetis, 2.5 mg/kg) was injected subcutaneously after the surgical procedures.

Visual stimulation. Visual stimuli were generated using the open-source Psychophysics Toolbox based on Matlab (MathWorks). Stimuli were presented at a distance of 15 cm to the left eye on a gamma-corrected LED-backlit LCD monitor (DELL) with a mean luminance of 20 cd/m^2. For experiments using a resonant scanner, the power source of the monitor's LED backlight was synchronized to the resonant scanner turnaround points (when data were not acquired) to minimize light leak from the monitor (Leinweber et al., 2014). We presented drifting sinusoidal gratings (2 Hz, 0.04 cycles/°, 100% contrast) unless stated otherwise. The trial structure of all stimulus sessions (receptive field mapping, orientation tuning, *et cetera*) was block randomized (the block size was given by the total number of parameter combinations).

Receptive field mapping: Stimuli consisted of a circular grating patch on a gray background (typically set to 20° in diameter). Stimuli were presented for 1 s at a single direction or for 2 s at the four cardinal

directions (0.5 s each). Stimulation periods were interleaved by 2 s of gray screen. We recorded 5 to 10 trials per stimulus condition.

Orientation tuning: We presented gratings of at least 10° diameter drifting in 8 directions (5 to 10 trials). Stimulus time was 2 s interleaved with 4 s of gray screen.

Size tuning: Patches of gratings were displayed at up to 9 different sizes, linearly spaced from 5° up to 85° in diameter (10 trials per size) centered on the ffRF. Stimulation time was 2 s interleaved by 4 s of gray screen. Stimuli were either presented at a single direction or at the four cardinal directions (0.5 s each).

Contextual modulation: We presented patches of gratings on a gray background (center stimulus), full-field gratings (iso stimulus), and patches of gratings on cross-oriented full-field gratings (cross stimulus). Stimulation time was 2 s interleaved by 4 s of gray screen. Trials with optogenetic stimulation had an additional 1 s pre-stimulus and post-stimulus gray screen during which the optogenetic light source was turned on and the total number of trials was doubled (**Optogenetics** below).

Behavioral monitoring. All mice were habituated (3 to 5 days) to experimental rigs before starting experiments. During all awake experiments, we recorded the positions of the left eye using a CMOS camera (DMK23UM021, Imaging Source) with a 50 mm lens (M5018-MP, Moritex), tracked the running speed of the mouse, and monitored its general behavior using a webcam (LifeCam Cinema 720p HD, Microsoft).

Two-photon calcium imaging. Imaging was performed using either a galvanometric-scanner based MOM (Sutter) or a resonant-scanner based (8 kHz) Bergamo II two-photon microscope (Thorlabs), both controlled by ScanImage (Vidrio). Using the MOM system, we acquired images of 128 × 128 pixels at a single depth at 5.92 Hz frame rate. With the Bergamo II, we acquired images of 380 × 512 pixels at 1 or 4 depths at 40 Hz or 8 Hz frame rate, respectively. We obtained similar results with both systems, so all data were pooled. The illumination light source was a Ti:sapphire laser (Chameleon Ultra II, Coherent) used at an excitation wavelength of 910 nm. The laser power under the objective (16×, Nikon) was typically set to 30 mW and never exceeded 50 mW (laser pulse width 140 fs at a repetition rate of 80 MHz).

Optogenetics. To silence VIP neurons, we used a 594 nm laser (OBIS 594 LS 100 mW, Coherent). We modified the Bergamo II microscope (Thorlabs) to combine optogenetic manipulation with two-photon calcium imaging. A lens (LA1805-B, Thorlabs) was placed in the optogenetic stimulation light path to defocus the light at the imaging plane. We used a dichroic mirror (DMBP740B, Thorlabs) to combine two-

photon laser and optogenetic stimulation light and a second one to split the green fluorescent protein (GFP) emission from both illumination light sources (FF555-Di03-25×36, Semrock). The laser for optogenetic stimulation was synchronized to the resonant scanner turnaround points (when data were not acquired) to minimize light leak from the monitor (Attinger et al., 2017; see **Visual Stimulation** for timing within a trial). The 594 nm laser power under the objective did not exceed 18 mW.

Data analysis. All data were analyzed using custom-written code in Matlab (MathWorks).

Two-photon calcium imaging: We analyzed two-photon calcium imaging data as described previously (Keller et al., 2020). Briefly, data were full-frame registered using custom-written software (https://sourceforge.net/projects/iris-scanning/). We selected the neurons semi manually, based on mean and maximum projection images. We calculated the raw fluorescence traces as the average fluorescence of all pixels within a selected region of interest for each frame. Fluorescence changes (ΔF/F) were calculated as described elsewhere (Dombeck et al., 2007). All stimulus evoked responses were baseline subtracted (1 s pre-stimulus interval).

Response amplitude: The response amplitude to a stimulus was computed as the average response over the duration of the stimulus presentation (excluding the first 0.5 s of each trial due to the delay and slow rise of calcium indicators). Responses were normalized by the maximum response over the relevant stimulus parameter space and then averaged over neurons. We defined significant responses as responses that exceeded a z-score of 3.29 (corresponding to $p < 10^{-3}$) or 4.41 (corresponding to $p < 10^{-5}$; for experiments in layer 4).

Receptive field mapping: To estimate the center of the receptive field, we fitted the responses to patches of gratings with a two-dimensional Gaussian. We excluded neurons if they failed to have at least one significant trial-averaged response within 10° of their estimated ffRF centers or if the ffRF centers were not within 10° of the centers of the stimuli presented to establish size tuning, orientation tuning, *et cetera*.

Size tuning: We fitted the integral over a difference of Gaussians. This fit was used to estimate the neurons' preferred sizes. We approximated the ffRF size by the size of the patch of gratings evoking the largest response (size tuning fits were bound to the interval 0.1 to 90.1°).

Orientation tuning: We fitted a circular sum of Gaussians with a peak offset of 180° and equal tuning width (full width at half maximum of the Gaussian fit). We calculated orientation selectivity index (OSI) as described elsewhere (Niell and Stryker, 2008).

Contextual modulation: To estimate the contextual modulation of excitatory, VIP, SOM, and PV neurons, we used a center patch diameter of 20°. We calculated a surround modulation index defined as the difference between the activity to cross and iso stimuli divided by the sum of the two. To estimate the effect of silencing VIP neurons on the contextual modulation of putative excitatory neurons, neurons were only considered if their preferred size was within 10° of the center patch diameter. Note that for these experiments, the center patch diameter was set to a size between 10° and 30°.

Statistics. We used Wilcoxon rank-sum tests for independent group comparisons, Wilcoxon signed-rank tests for paired tests and Student's t-tests for a single group analysis. No statistical methods were used to pre-determine sample sizes, but our sample sizes were similar to those used in previous publications.

Data availability. Data sets supporting the findings of this paper are available on request from the corresponding authors.

Code availability. Custom code is available from the corresponding authors on request.