



Published in final edited form as:

Nat Neurosci. 2020 January ; 23(1): 61–74. doi:10.1038/s41593-019-0552-7.

Prefrontal somatostatin interneurons encode fear memory

Kirstie A. Cummings¹, Roger L. Clem^{1,*}

¹Department of Neuroscience and the Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY

Abstract

Theories stipulate that memories are encoded within networks of cortical projection neurons (PNs). Conversely, GABAergic interneurons (INs) are thought to function primarily to inhibit PNs and thereby impose network gain control, an important but purely modulatory role. Here we show in male mice that associative fear learning potentiates synaptic transmission and cue-specific activity of medial prefrontal cortex (mPFC) somatostatin interneurons (SST-INs), and that activation of these cells controls both memory encoding and expression. Furthermore, the synaptic organization of SST- and parvalbumin (PV)-INs provides a potential circuit basis for SST-IN-evoked disinhibition of mPFC output neurons and recruitment of remote brain regions associated with defensive behavior. These data suggest that rather than constrain mnemonic processing, potentiation of SST-IN activity represents an important causal mechanism for conditioned fear.

Associative memory is a critical function of cortical brain networks, which are primarily populated by excitatory PNs and inhibitory INs. The most abundant of these cell types, PNs are a key substrate for interregional brain signaling that is critical for memory expression^{1,2}. Accordingly, retrieval cues activate subsets of PNs that are hypothesized to encode stimulus associations through persistent changes in excitatory synapse strength and density³. In contrast, GABAergic INs are generally thought to inhibit PNs^{4–8}, which has been suggested to play a role in optimizing the dynamic range of PN firing to indirectly modulate the strength and specificity of learning. However, while several studies credit INs with such “supporting roles”, it remains unclear whether they can directly mediate the encoding of cue associations through their own functional plasticity.

Fear conditioning is a powerful model of such learning in which an animal acquires survival-based defensive reactions to a conditioned stimulus (CS) that predicts imminent threat. The expression of fear memory in rodents requires neural activity in the prefrontal subregion of

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence should be addressed to: roger.clem@mssm.edu.

Author contributions

K.A.C. and R.L.C. initiated the project; R.L.C. supervised research; K.A.C. and R.L.C. designed experiments; K.A.C. performed the research and data analysis; R.L.C. and K.A.C. wrote the manuscript.

Competing interests

We declare no conflicts of interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

mPFC⁹, where both PNs and INs sampled by extracellular recordings exhibit CS-evoked changes in firing rate after conditioning⁴. However, it is unknown whether learning induces synaptic plasticity in prelimbic circuits and, if so, whether IN activity is modulated by these changes in conjunction with memory encoding. Here we address these questions in mice with a combination of synaptic electrophysiology, calcium imaging, optogenetic manipulation and brain activity mapping of prelimbic interneurons and associated circuitry. We demonstrate that SST-INs exhibit properties indicative of a memory storage substrate, including **1.** learning-dependent potentiation of synaptic transmission, **2.** cue-specific activation during memory retrieval, and **3.** bidirectional modulation of memory expression. Moreover, prelimbic SST-INs exert potent disinhibitory control over a fear-related brain network, suggesting a fundamental role for these cells in orchestrating conditioned fear responses.

Results

Cued fear learning potentiates SST-IN excitatory input

While experience-dependent plasticity is considered to be the most likely mechanism for cortical information storage^{2,3}, the extent to which learning is associated with plasticity of cortical inhibitory circuits remains poorly understood. To determine whether fear learning alters the synaptic properties of prefrontal INs, we therefore obtained *ex vivo* electrophysiological recordings from parvalbumin (PV)- and SST-expressing cells, which together comprise the majority of cortical GABAergic interneurons¹⁰. In order to identify these cell types in acute brain slices we generated IN-specific expression of tdTomato by crossing the Ai9 reporter line to PV- or SST-Cre driver mice, which exhibit highly selective recombination in prelimbic cortex (Fig. 1a)¹¹. We then subjected these animals to behavioral training entailing either paired or unpaired presentations of an auditory CS and footshock (unconditioned stimulus, US; Fig. 1b; Supplementary Fig. 1a, f). Because only paired animals acquire CS-evoked defensive freezing, unpaired mice served as a control for non-associative effects of stimulus exposure^{12,13}. At 24 hours after training, spontaneous excitatory (sEPSCs) and inhibitory postsynaptic currents (sIPSCs) were recorded and analyzed as a proxy for potential synaptic plasticity in prelimbic INs (Fig. 1b–c; Supplementary Fig. 1). Mice that received CS-US pairing displayed higher sEPSC frequency in SST- but not PV-INs residing in layer 2/3 (L2/3), compared to naïve and unpaired controls (Fig. 1c). No other differences in sEPSC or sIPSC properties were associated specifically with CS-US pairing (Supplementary Fig. 1). Because sEPSC frequency can reflect differences in presynaptic efficacy, we next measured the response of layer 2/3 SST-INs to local paired-pulse stimulation, a well-established assay for neurotransmitter release probability (Fig. 1d)¹⁴. Consistent with increased glutamate release onto SST-INs after fear conditioning, evoked EPSCs exhibited a higher paired-pulse ratio in animals that received CS-US pairing but not unpaired training. These results confirm that cued fear learning is associated with potentiation of excitatory synapses onto SST-INs.

SST-INs signal the conditioned stimulus after learning

Given the observed learning-related potentiation of their excitatory input (Fig. 1), we hypothesized that SST-INs in conditioned mice might exhibit increased CS-related activity

and thereby participate in memory processing. To test this possibility, we utilized fiber photometry to monitor *in vivo* activity of SST-INs after conditional viral expression of GCaMP6f (AAV-DIO-GCaMP6f-eYFP) in SST-Cre mice (Fig. 2; Supplementary Fig. 2). Mice were unilaterally implanted with an optic fiber (400 μm core diameter) directed at the prelimbic cortex and imaged under freely-behaving conditions during CS exposure before, during, and after auditory fear conditioning (Fig. 2b; Supplementary Fig. 3). During CS-US pairing, we observed an increase in CS-related calcium signals during trials 5–6 compared to the initial conditioning trial (Fig. 2b–c), which mirrored a trial-dependent increase in CS-evoked freezing (Supplementary Fig. 3). One day after conditioning, CS-related calcium signals during re-exposure to the CS were markedly increased compared to a preconditioning test (Fig. 2b,d). Because CS presentation leads to defensive freezing, the recruitment of SST-INs during memory retrieval could be a consequence of fear expression rather than CS modulation per se. We therefore additionally analyzed calcium signals during inter-trial freezing bouts, during which fear-related SST-IN activity can be dissociated from CS processing. Unlike CS trials, transitions from movement to freezing during the inter-trial intervals were associated with negligible fluorescence changes (Fig. 2d).

To further test whether fearful states modulate SST-IN activity independent of the CS, we next performed SST-IN-specific fiber photometry in mice that were subjected to unpaired CS-US training (Supplementary Fig. 4a–c). Importantly, although unpaired training results in context conditioned fear, it does not induce synaptic plasticity in SST-INs (Fig. 1c–d). During unpaired conditioning, US trials were associated with large calcium signals, confirming that changes in SST-IN activity could be readily detected (Supplementary Fig. 4d, f). However, regardless of whether imaging was conducted in a novel context (context B) or the original training arena (context A), no changes in calcium signals were associated with spontaneous freezing bouts (Supplementary Fig. 4e–f). Moreover, there were no overall differences in the frequency of calcium transients (peaks) during exposure to contexts A and B, relative to a preconditioning baseline in context A (Supplementary Fig. 4g), despite robust differences in freezing between these tests (Supplementary Fig. 4b). These data indicate that SST-INs do not generally signal a high-fear state, but are instead specifically activated by the threat-associated cue.

To independently confirm the activation of SST-INs in response to memory retrieval, we performed immunohistochemical labeling for the activity reporter cFos (Supplementary Fig. 5), which is considered to be a marker of neurons strongly activated by mnemonic cues³. To elicit memory retrieval in SST-Cre/ Ai9 reporter mice, 4 CS trials were presented at 24 hours after fear conditioning in a context distinct from the training arena, and as control conditions we examined mice in which either conditioning or memory retrieval were omitted. Retrieval CS trials triggered a robust increase in freezing only in mice that had been previously conditioned (Supplementary Fig. 5b–d). Following behavioral testing, substantially more SST-INs exhibited cFos immunoreactivity in these animals compared to those in either control group (Supplementary Fig. 5a, e–f). Increased SST-IN activation in conditioned mice could also be observed specifically in layer 2/3, consistent with a causal role for lamina-specific SST-IN plasticity (Fig. 1). In contrast to these results, PV-INs did not exhibit a detectable increase in cFos immunoreactivity under the same conditions in PV-Cre/ Ai9 mice (Supplementary Fig. 6).

SST-IN activity controls memory expression

CS activation of SST-INs suggests that recruitment of these cells could be important for memory retrieval. On the other hand, previous studies suggest that SST-IN activity might also function to constrain fear expression through inhibition of cue-responsive PNs⁴⁻⁸. To establish the behavioral impact of SST-IN recruitment, we utilized optogenetics to modulate SST-IN activity in conjunction with CS-evoked memory retrieval. We microinjected into prelimbic cortex viral vectors encoding archaerhodopsin (Arch; AAV-FLEX-Arch3.0-GFP), channelrhodopsin-2 (ChR2; AAV-DIO-ChR2-eYFP), or an opsin-negative control vector (eYFP; AAV-DIO-eYFP). Importantly, optic illumination was sufficient to hyperpolarize or depolarize Arch- or ChR2-expressing SST-INs, respectively, and thereby reliably control the firing of these cells (Supplementary Fig. 7a–b). Following virus infusion, mice were implanted with optic fibers directed at the prelimbic cortex (Fig. 3a; Supplementary Fig. 7c–d). One week after surgery, these mice were subjected to CS-US pairing in the absence of photostimulation.

At 24 hours after training, a memory retrieval test was conducted in which the independent and combined effects of light and CS were examined in a context distinct from the training arena. Compared to CS-only trials, a marked reduction in freezing was observed in Arch-expressing mice during light (532 nm, 20 s, constant) + CS trials (Fig. 4b,d), indicating that SST-IN activity is required for cued memory expression. No reduction in freezing was observed in Arch animals during light-only trials compared to the baseline period, suggesting that SST-IN activity is not required for generalized context fear. Conversely, in ChR2-expressing mice, photoexcitation of SST-INs (473 nm, 10 ms pulses, 20 Hz) even in the absence of CS presentation was sufficient to increase freezing over baseline levels and thereby mimic a conditioned response (Fig. 3c,d). Combined CS + light presentation in these animals did not increase freezing beyond that observed during interleaved CS-only trials, suggesting the possibility of a ceiling effect. Importantly, light-evoked freezing cannot be explained by a nonspecific motor deficit because photoexcitation of ChR2-expressing SST-INs did not alter standard metrics of locomotion in the open field test (Supplementary Fig. 8). Finally, no light effects were observed in eYFP control groups that were stimulated with the same parameters used in Arch (Fig. 3b,d) or ChR2 mice (Fig. 3c–d), and all groups displayed a CS-evoked increase in freezing of similar magnitude and therefore did not differ in memory strength (Fig. 3e).

To establish whether fear-promoting properties are intrinsic to SST-INs or acquired through learning, we performed photoexcitation of SST-INs without prior fear conditioning. In contrast to fear conditioned mice (Fig. 3c), no consistent effect of photoexcitation was observed in naïve subjects (Fig. 3f).

SST-IN activation and plasticity mediates memory acquisition

Given the critical role of SST-INs in memory expression, we next sought to utilize photoinhibition to determine whether SST-IN activity at the time of CS-US pairing is required for memory acquisition as well as associated plasticity of SST-INs. After prelimbic infusion of AAV-FLEX-Arch3.0-GFP or eYFP control vectors, SST-Cre mice were subjected to CS-US pairing, during which photoinhibition was timed to coincide with each

of the 6 CS-US trials (Fig. 4a; Supplementary Fig. 9). On the following day, memory retrieval was examined in the absence of photoinhibition during presentation of 4 CS retrieval trials. Remarkably, while CS presentations triggered an increase in freezing in both Arch-expressing mice as well as eYFP controls, the magnitude of CS-evoked increase in freezing was dramatically lower in Arch-expressing mice (Fig. 4b–c). This effect cannot be attributed solely to Arch expression, because a retrieval deficit was not observed in Arch-expressing mice that were conditioned without photoinhibition (Fig. e). Following memory retrieval, acute brain slices were obtained to compare excitatory synaptic transmission in prelimbic SST-INs of Arch versus eYFP subjects, focusing on layer 2/3 cells located within 1 mm below the optic fiber track. Compared to eYFP controls, Arch-expressing SST-INs exhibited lower sEPSC frequency as well as higher paired-pulse ratios of evoked EPSCs (Fig. 4d–e).

While these results suggest that prefrontal SST-INs participate in memory encoding, previous studies have established that fear conditioning requires basolateral amygdala (BLA) activity¹⁵, and memory storage is widely believed to be mediated by plasticity of synaptic connections within the lateral and basal nuclei^{12,13,16–22}. Therefore, because activation of prelimbic SST-INs evokes a fear response (Fig. 3c), we wondered whether deficits in learning after amygdala silencing can be explained in part by reduced prefrontal SST-IN transmission. To test this hypothesis, we performed photoinhibition of BLA excitatory neurons during CS-US pairing. One week prior to conditioning, SST-Cre x Ai9 mice received BLA injections of Arch or eYFP control vectors under the control of a CaMKII promoter (AAV-CaMKII-ArchT-GFP or AAV-CaMKII-eYFP) and were implanted with optic ferrules directed at the BLA (Fig. 4f; Supplementary Fig. 10). Fear conditioning and retrieval tests were conducted as described for prelimbic SST-IN photoinhibition (Fig. 4f). While freezing occurred in both groups during conditioning (Fig. 4g), Arch subjects exhibited a deficit in CS-evoked responses relative to eYFP subjects during the memory retrieval test (Fig. 4h). Following retrieval, whole-cell recordings revealed that SST-INs exhibited lower spontaneous EPSC frequency as well as increased paired-pulse ratios of evoked EPSCs (Fig. 4i–j) in Arch-expressing mice, compared to eYFP controls. These data independently confirm the relationship between memory encoding and increased SST-IN transmission, and suggest that SST-IN plasticity depends at least in part on BLA activity.

Microcircuit organization and opposing behavioral roles of SST- and PV-INs

An increase in freezing upon SST-IN photoexcitation (Fig. 3c) is surprising given that SST-INs are likely to form inhibitory contacts onto excitatory PNs that control memory expression¹⁵. However, we considered the possibility that another class of interneurons also receives input from SST-INs and, when inhibited via these connections, might be responsible for disinhibition of PNs²³. In particular, previous work indicates that mPFC PV-INs exhibit firing rate inhibition during CS exposure, and that photoinhibition of PV-INs elicits freezing⁴. We therefore utilized optogenetics to establish whether SST-INs form functional connections with PV-INs, as well as to compare the strength of such connections with those formed onto PNs from the same brain slices. In addition, we asked whether CS-US pairing affects the balance of transmission from SST- or PV-INs onto inhibitory versus excitatory

populations, which might modulate whether SST- or PV-IN recruitment leads predominantly to inhibition or disinhibition of PNs.

First, to enable selective interrogation of synaptic outputs from SST-INs, we microinjected an INTRSECT ChR2 vector (AAV-Cre_{off}/Flp_{on}-hChR2-eYFP)²⁴ into the prelimbic cortex of SST-IRES-FlpO/PV-IRES-Cre/Ai9 triple transgenic mice, permitting independent tagging of SST- (ChR2) and PV-INs (tdTomato) (Fig. 5a). We then examined inhibitory responses to SST-IN photoexcitation in PV-INs as well as PNs from the same brain slices to control for group differences attributable to viral expression. In both naïve and unpaired control mice, SST-INs elicited monosynaptic IPSCs in PV-INs that were less potent than those recorded in surrounding PNs (Fig. 5b–c; Supplementary Fig. 11). In contrast, after CS-US pairing, responses to SST-IN photoexcitation in these cell types were similar in amplitude.

Comparison of SST-IN-evoked responses in PV-INs normalized to those in PNs confirmed that CS-US pairing increases the relative strength of SST-IN→PV-IN transmission (Fig. 5d). This suggests that learning shifts the balance of SST-IN output to favor inhibition of PV-INs, which may increase the likelihood of SST-IN-evoked disinhibition.

Next, we utilized a similar genetic approach to interrogate PV-IN transmission onto SST-INs and PNs in prelimbic layer 2/3 (Fig. 5f). Strikingly, this revealed that regardless of the training condition, PV-INs elicit IPSCs that are ~10-fold larger in amplitude in PNs compared to surrounding SST-INs (Fig. 5g–h). Comparison of SST-IN responses normalized to those in PNs revealed no effect of training on the balance of transmission (Fig. 5i). These results indicate the presence of a strong bias in PV-IN output that potentially favors preferential control of PN over SST-IN firing (Fig. 5j). In contrast, SST-INs exhibit a much weaker bias (~2-fold) for PNs over PV-INs, and this bias is completely eliminated by conditioning (Fig. 5e). This implies that SST-INs in the prelimbic cortex could have a unique capacity to evoke PN disinhibition.

Prelimbic circuit organization suggests that SST-INs might interact directly with PV-INs to mediate fear expression through PN disinhibition. As an *in vivo* test of this model, we first sought to confirm that photoinhibition of PV-INs elicits freezing, as reported previously⁴. Indeed, after fear conditioning, PV-IN photoinhibition resulted in an increase in freezing during light-only trials (Fig. 6a, Supplementary Fig. 12). We then tested whether inhibition of PV-INs is required specifically for SST-IN-evoked freezing. In contrast to SST-IN-specific manipulations (Fig. 3c), concurrent photoactivation of SST- and PV-INs in SST-Cre/PV-Cre double transgenic mice negated the fear-promoting effect of SST-IN activity (Fig. 6b). These results imply that interaction between SST- and PV-INs is important for processing the behavioral output of SST-IN activity.

BLA afferent connectivity of prelimbic INs

Having established that SST-INs directly inhibit PV-INs, we next considered whether these IN populations are engaged by long-range inputs to the prelimbic cortex. In addition to modulating plasticity of prelimbic SST-INs (Fig. 4), BLA contains prelimbic cortex-projecting neurons that exhibit increased firing during memory retrieval^{25,26} and regulate fear memory expression^{27,28}. Projections from these cells primarily target layer 2/3 and are thus well-positioned to recruit potentiated SST-INs²⁹. To test whether prelimbic interneurons

are directly modulated by these projections, we infused a calcium-calmodulin kinase II (CaMKII)- driven ChR2 vector (AAV-CaMKII-hChR2-eYFP; CaMKII-hChR2) into BLA (Fig. 7a), leading to axonal ChR2 accumulation in prelimbic layer 2/3 of SST- (Fig. 7b) and PV- and -IRES-Cre/ Ai9 mice (Fig. 7g). Optic stimulation of these projections in naïve mice elicited compound EPSCs and feedforward IPSCs in both SST- and PV-INs, as well as surrounding PNs (Supplementary Fig. 13). Because much of this transmission occurred at long latencies after stimulation, recurrent activity within prelimbic circuits is likely responsible for its generation. Interestingly, compared to PNs from the same brain slices, SST- but not PV-INs exhibited a higher ratio of excitatory to inhibitory charge during these events (Supplementary Fig. 13c). This could be attributed to less potent network inhibition of SST-INs, since inhibitory charge in SST-INs was lower than in surrounding PNs (Supplementary Fig. 13b).

While the above results are intriguing, the complexity of BLA-evoked activity prohibited analysis of monosynaptic transmission at connections between BLA axons and prelimbic INs. We therefore applied a pharmacological cocktail to eliminate action potential propagation and prevent polysynaptic transmission^{12,13,30}. When using this approach in SST-Cre/ Ai9 mice, we found that regardless of training condition, BLA afferents evoke responses of similar amplitude in SST-INs compared to PNs (Fig. 7c–d; Supplementary Fig. 14). However, when SST-IN responses were normalized to those of PNs, this revealed a slightly higher ratio of SST-IN/ PN transmission in paired mice compared to naïve controls (Fig. 7e). In contrast to the above results, BLA terminal stimulation in naïve and unpaired PV-Cre/ Ai9 mice evoked EPSCs that were larger in amplitude in PV-INs compared to PNs (Fig. 7h–j). This is in agreement with similar experiments that examined the potency of BLA transmission onto PV-INs in the infralimbic cortex³¹. However, in animals that received CS-US pairing the relative strength of BLA transmission in PV-INs and PNs was effectively reversed (Fig. 7j). These data collectively imply that, following conditioning, SST-INs are likely strongly activated by BLA afferents and that circuit plasticity may favor their recruitment over PV-INs (Fig. 7f, k).

Network disinhibition underlies SST-IN-evoked fear expression

Because the balance of ongoing excitatory and inhibitory transmission determines the firing rate of excitatory PNs, the activity of prelimbic output neurons could be modulated solely through the relief of PV-IN-mediated inhibition. To reveal the extent to which SST-IN activity disinhibits prelimbic networks, we therefore conducted immunolabeling for the activity marker cFos in prelimbic cortex and potential downstream brain regions following SST-IN photoexcitation, in the absence of any CS exposure, at 24 hours after conditioning (Fig. 8a; Supplementary Fig. 15). Similar to a previous experiment (Fig. 3), ChR2-expressing mice but not eYFP controls exhibited an increase in freezing in response to photoexcitation (Supplementary Fig. 15). After behavioral testing, ChR2 mice exhibited higher cFos labeling of SST-INs compared to eYFP controls, as well as higher cFos labeling of surrounding eYFP-negative cells, consistent with disinhibition of other prelimbic cell types (Fig. 8b, d). When quantification was extended to downstream targets of prelimbic cortex, higher numbers of cFos-positive cells were also detected in the BLA, paraventricular thalamus, lateral habenula, ventrolateral periaqueductal gray, and the dorsomedial

hypothalamus (Fig. 8c, e). However, several other regions including the nucleus accumbens, caudate putamen, ventral hippocampus area CA1, dentate gyrus, and mediodorsal thalamus were unaffected by photostimulation (Fig. 8e; Supplementary Fig. 15).

To test whether regional cFos induction by SST-IN photoexcitation occurs independently of fear expression, we next quantified cFos expression in a randomly selected subset of naïve mice that received optogenetic manipulation of prelimbic SST-INs without prior fear conditioning (Fig. 3f). Consistent with the larger group, these mice exhibited no increase in freezing over baseline levels during photostimulation (Fig. 8f; Supplementary Fig. 16). Examination of stimulated prelimbic tissue confirmed that consistent with photoactivation, higher cFos labeling was present in eYFP-positive cells of ChR2- relative to eYFP-expressing mice (Fig. 8g; Supplementary Fig. 16). However, there was no group difference in cFos expression in surrounding eYFP-negative cells, indicating that in contrast to animals that received CS-US pairing (Fig. 8b, d), SST-IN photoexcitation in naïve mice does not activate surrounding prelimbic neurons to a significant degree. In addition, remote brain regions that were modulated by photoexcitation in conditioned mice did not exhibit any group differences in the number of cFos-positive neurons (Fig. 8h; Supplementary Fig. 16). Thus, acquisition of SST-IN-evoked freezing correlates with a change in SST-IN recruitment of a specific brain network including prelimbic neurons indirectly activated by SST-IN photoexcitation, presumably via disinhibition.

Finally, to test whether network-level effects of memory retrieval resemble those evoked by SST-IN photoexcitation in conditioned mice, we performed cFos analysis following CS exposure (Fig. 8i). Presentation of 4 CS trials elicited increased freezing in mice that received CS-US pairing 24 hours prior to the memory retrieval test, but not in non-conditioned controls (Supplementary Fig. 17). After CS exposure, a higher number of cFos-positive cells was observed in conditioned relative to non-conditioned mice in the majority (5/6) of brain regions that were modulated by SST-IN photoexcitation (Fig. 8j). The remaining region (vIPAG) showed a trend toward higher cFos labeling in conditioned mice ($p = 0.082$). Conversely, areas in which cFos immunoreactivity was unaffected by SST-IN photoexcitation also exhibited no differences in cFos-positive cells following CS-evoked memory retrieval. Together these data argue against the notion that network cFos induction by photostimulation results from nonphysiological activity patterns, and suggest that CS recruitment of SST-INs mediates disinhibition of prelimbic outputs to remote brain regions underlying memory expression.

Discussion

In this study we demonstrate that associative fear conditioning potentiates the function of prefrontal SST-INs at the level of synaptic transmission and *in vivo* activity. Correlated with memory acquisition, SST-INs exhibit an experience-dependent increase in CS-evoked firing and specifically signal the learned CS rather than the general fear state of the animal. Activation of SST-INs in turn plays a causal role in both the expression as well as initial acquisition of cued fear responses. The paradoxical role of SST-INs in memory expression can be explained by their encoding of cue-related disinhibition of prelimbic PNs, which are in turn responsible for the recruitment of a distributed brain network for defensive

responding. Although our results do not exclude the possibility of similar role in context-evoked fear, SST-IN activity was not modulated by the conditioning context (Supplementary Fig. 4) and there was no baseline effect of SST-IN photoinhibition on generalized context freezing (Fig. 3), suggesting that SST-INs preferentially mediate the encoding of cue associations.

In network disinhibition, prelimbic circuits formed by SST- and PV-INs exhibit functional differences that may facilitate their complementary roles. In particular, output from PV-INs is heavily biased toward PNs over SST-INs (Fig. 5), implying that they are specialized for suppression of PN firing. In contrast, there is a greater potential for PN disinhibition during SST-IN activity, owing to their relatively strong inhibition of PV-INs. Accordingly, photoactivation of SST-INs indirectly activates not only surrounding prelimbic cells but also remote brain regions that receive excitatory connections from the mPFC (Fig. 8). A learning-dependent shift in SST-IN output could in part explain why these effects are specific to conditioned mice. Alternatively, plasticity in other circuits could prime this network for SST-IN-evoked disinhibition.

Although all prelimbic cell types that we examined receive direct input from BLA projections, it is notable that these afferents evoke complex synaptic activity with an overall higher ratio of excitatory: inhibitory transmission in SST-INs (Supplementary Fig. 13). In conditioned mice, PV-INs but not SST-INs exhibited a lower strength of monosynaptic BLA transmission compared to neighboring PNs (Fig. 7). Given these findings, it seems possible that after learning BLA afferent activity favors the recruitment of SST-INs over PV-INs, and that resulting PN disinhibition could underlie freezing behavior elicited by BLA afferent stimulation in conditioned mice²⁷. In addition, BLA projections could also mediate the effect of amygdala activity on learning-related SST-IN plasticity (Fig. 4). Further interrogation of synaptic and functional plasticity throughout this circuitry will be necessary to establish the degree to which other cell types, and projection-specific populations, cooperate with SST-INs in memory encoding and retrieval.

An important implication of SST-IN plasticity is that phasic disinhibition, which has been established primarily as a mechanism for cortical processing of behavioral feedback signals^{23,32–35}, can be encoded by potentiation of GABAergic transmission to mediate novel cue associations. This means that both activated (e.g. SST-INs) and inhibited interneuron populations (e.g. PV-INs) can play a role in memory expression and exhibit cue-evoked activity patterns that are causally linked. In this way the modification of a sparse interneuron population can extensively reorganize stimulus processing in a broader local network. Indeed, previous findings confirm that firing rate inhibition of PV-INs is a potent regulator of PN recruitment, synchrony, and synaptic plasticity^{4,8,23,32,35}. Nevertheless, it remains unclear why SST-IN-evoked disinhibition triggers such selective changes in behavior and circuit activity. One possibility is that different subnetworks of prelimbic PNs, for example populations with discrete projection patterns, vary in the level of inhibition they receive from SST- versus PV-INs. The balance of ongoing activity from these inhibitory sources could in part determine whether a given cell is predominantly activated, inhibited or unaffected by cue-responsive SST-INs.

Several brain regions activated by prelimbic SST-IN stimulation, including the BLA¹⁵, paraventricular thalamus^{36,37}, lateral habenula³⁸, ventrolateral periaqueductal gray³⁹, and the dorsomedial hypothalamus⁴⁰, have established roles in mediating defensive behavioral responses to conditioned and unconditioned stimuli. It is possible that PNs that give rise to direct projections to these regions from the prelimbic cortex are preferentially disinhibited during SST-IN firing. Alternatively, this pattern of network activity may be completed by relay circuits from other downstream effectors. Intriguingly, however, neither the striatum nor the mediodorsal thalamus, which are among the areas receiving the highest density of input from the prelimbic cortex^{41,42}, are activated during the expression of SST-IN-evoked freezing. This implies a high level of specificity in the computational logic of prelimbic microcircuits.

While our *in vivo* manipulations were targeted to SST-INs as a whole, somatostatin expression demarcates a cell population with heterogeneous protein expression, firing patterns and morphology⁴³, whose function is additionally determined by their laminar location⁴⁴. Our results imply that SST-INs that encode fear-related disinhibition reside in superficial layers of the prelimbic cortex, but further parcellation of this cell class could reveal functionally discrete subpopulations. For example, previous work has identified a subset of SST-INs that express the oxytocin receptor and contribute to the regulation of anxiety and social behavior in a sex-dependent manner^{45,46}. At present it remains unclear what features of an aversive experience might be encoded by SST-INs or how this information is represented by specific SST-IN populations. Future studies could resolve the function of individual neurons through the use of genetic markers, cellular tags and *in vivo* imaging. In addition, although circuit disinhibition can be explained by fast GABAergic transmission, it is also important to consider that somatostatin is not just a marker of interneurons but also a peptide transmitter that can influence memory acquisition⁴⁷ and recall⁴⁸. The release of this peptide from dense-core vesicles may contribute to the memory function of potentiated SST-INs.

In conclusion, our results outline an important casual role for inhibitory signaling in associative memory. In pursuit of memory engrams, it will therefore be critical to consider the contributions of interneurons that, despite their inhibitory output, have the capacity to encode and reactivate a specific pattern of excitatory neuronal activity. Interrogation of these cells and their associated circuitry could reveal important computational principles for the storage, consolidation and retrieval of information. Indeed, during the preparation of our revised manuscript, a new report demonstrated an important causal role for prefrontal SST-IN activity in learned social fear⁴⁹.

Methods

Animals

All experimental procedures were approved by the Institutional Animal Care and Use Committee at the Icahn School of Medicine at Mount Sinai. Experiments were performed on male mice aged P42-P60. Mice were acquired from Jackson Laboratories (Bar Harbor, ME, USA) and maintained in the C57Bl/6J background. The following genotypes were used: SST-IRES-Cre (Stock No. 028864), PV-IRES-Cre (Stock No. 017320), SST-IRES-FlpO

(Stock No. 028579), PV-IRES-FlpO (Stock No. 022730), and Ai9 (Stock No. 007909). Mice were housed 2–5 per cage in a 12 hr light-dark cycle with access to food and water *ad libitum*.

Stereotaxic vector infusion and optic fiber implantation for *in vivo* optogenetics

Viral constructs were purchased from University of Pennsylvania Vector Core or Addgene and included AAV1-EF1a-DIO-hChR2(H134R)-eYFP-WPRE (Addgene # 20298), AAV1-CBA-FLEX-Arch-GFP (Addgene # 22222), AAV1-EF1a-DIO-eYFP-WPRE (Addgene # 27056), AAV1-CaMKII α -ArchT-GFP (Addgene # 99039), and AAV1-CaMKII α -eYFP-WPRE (Addgene # 105622). Optic fibers (200 μ m diameter, Thor labs) were fixed inside ferrules (1.25 mm OD, 230 μ m ID, Precision Fiber Products, Chula Vista, CA, USA) using heat-cured epoxy, cut to 2.5 mm length for mPFC or 5.0 mm length for basolateral amygdala, and polished using aluminum oxide lapping paper with incremental decreases in graininess (Thor labs, Newton, NJ, USA). Light transmittance was measured using a light intensity meter (Thor labs) at wavelengths used in each respective experiment. Stereotaxic surgeries were conducted at P45–49. Anesthesia was induced using 5% inhaled isoflurane vaporized in oxygen delivered at a rate of 1–1.5 L/min. Mice were then mounted in stereotaxic frames and maintained at 1–1.5% isoflurane (Stoelting, Wood Dale, IL, USA or Kopf, Tujunga, CA). Viral constructs were injected bilaterally into prelimbic cortex (300 nL; AP +1.9, DV –2.0, ML +/- 0.9 at a 10° angle) or basolateral amygdala (250 nL; AP –1.4, DV –5.1, ML +/- 3.3) using motorized injectors (Stoelting and World Precision Instruments) at a rate of 100 nL/min. Following infusion, the needle was left in place for an additional 10 minutes prior to slow removal at a rate of 0.03 mm/second. For *in vivo* optogenetics, following bilateral injections, optic ferrules were bilaterally implanted and directed towards prelimbic cortex (AP +1.9, DV –1.6, ML +/- 0.9 at a 10° angle) or basolateral amygdala (AP –1.4, DV –4.8, ML +/- 3.3). Ferrules were fixed in place using C&B Metabond luting cement (Parkell, Edgewood, NY, USA) and dental cement. Post-surgical analgesia was achieved with banamine (2.5 mg/kg). Mice recovered in home cages for at least 1 week prior to experimental manipulation.

Stereotaxic vector infusion for electrophysiology

AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE (Addgene # 26969P) was purchased from the University of Pennsylvania Vector Core. rAAVDJ/nEF-Cre_{off}/Flp_{on}-hChR2(H134R)-eYFP was purchased from the University of North Carolina Gene Therapy Center Vector Core. Stereotaxic surgeries were done at P42–45 for AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE and from P25-P28 for AAVDJ-Cre_{off}/Flp_{on}-hChR2(H134R)-eYFP. AAV1-CaMKII α -hChR2(H134R)-eYFP-WPRE and rAAVDJ/nEF-Cre_{off}/Flp_{on}-hChR2(H134R)-eYFP were bilaterally infused into basolateral amygdala (250 nL; AP –1.3, DV –5.1, ML +/- 3.3) or prelimbic cortex (400 nL; AP +1.9, DV –1.6, ML +/- 0.2), respectively. Mice recovered for 10 days (AAV1-CaMKII α -eYFP-WPRE) or 25–27 days (rAAVDJ/nEF-Cre_{off}/Flp_{on}-hChR2(H134R)-eYFP) prior to electrophysiological recordings, which we determined to be the incubation time required for effective expression of ChR2 by these viral serotypes.

Optogenetic behavioral manipulations

Mice were acclimated to handling and patch cord tethering for 3 consecutive days prior to behavioral testing. Handling consisted of 10 min handling sessions, followed by 15 min habituation to patch cords in a fresh cage, followed by 5 min of additional handling. Fear conditioning and retrieval tests were conducted in sound attenuating chambers with automated stimulus delivery software (MedAssociates, St. Albans, VT, USA). Each session began with a 200 s baseline period after which various stimuli were presented with a fixed 80 s interstimulus interval. Auditory fear conditioning entailed 6 pairings of an auditory tone (CS; 2kHz, 80 dB, 20 s) with a co-terminating footshock (US; 0.7 mA, 2 s). For experiments where optogenetic manipulations were performed during memory retrieval, animals were conditioned while connected to patch cords but no light stimuli were delivered. Twenty-four hours after conditioning, modulation of freezing behavior by CS and light stimuli were examined in a context distinct from the conditioning arena (context B). In the retrieval test, mice were subjected to two laser-only trials, followed by 4 CS presentations alternating with and without concurrent laser stimulation. CS-only or CS + light presentation order was counterbalanced. The two laser-only trials, two CS-only, and two CS+light trials were each averaged and used for subsequent analysis. In experiments in naïve mice, animals were subjected to two trials of laser stimulation in context B without prior fear conditioning. For trials including optogenetic stimulation, a final transmitted intensity of 7–9 mW for 473 nm (for ChR2; 20 Hz, 10 ms for 20 s epochs) or 563 nm (for Arch; constant light, 20 s) laser-generated (Opto Engine, LLC, Midvale, UT, USA) light was used. Parameters for photoexcitation are consistent with other recent studies of cortical SST-INs^{50,51}, and the 20 Hz firing rate observed during *ex vivo* stimulation of SST-INs (Supplementary Fig. 7) is within the endogenous range of activity in prelimbic SST-INs⁵². In experiments where neurons were silenced during learning, light was presented simultaneously with each CS and extended 3 seconds after CS/US termination, after which it was manually ramped down to 0 mW over a period of 3 seconds. Twenty-four hours later, mice were tethered to patch cords in context B and subjected to 4 CS presentations in the absence of light stimulation. Mice were then used for electrophysiology recordings 24 hours after retrieval. Behavior was recorded by video and scored offline by a trained experimenter blind to the identity of the subject. Scoring by a second blinded experimenter was used to validate results. Fiber tip placement and viral expression were confirmed by fluorescence microscopy. Subjects having misplaced fiber tips or lacking viral expression encompassing at least ~50% of the target structure, in either hemisphere, were excluded from analysis by an investigator blind to experimental results.

Open field

Naïve SST-IRES-Cre mice expressing ChR2 or eYFP in prefrontal SST-INs were used for open field experiments. Mice were habituated to the room for 30 minutes prior to the start of the experiment. Mice were tethered to patch cords and placed into the center of a 42 cm (length) × 42 cm (width) × 30 cm (height) square arena equipped with 15 infrared beams/detectors on each side. Tests were 20 min long and consisted of alternating light on (473 nm, 7–9 mW delivered at 20 Hz, 10 ms pulses) and light off periods lasting 5 minutes each. Light on/off epochs were counterbalanced in both groups. Open field arenas were connected to a PC running Fusion v5.6 SuperFlex software, which was used to analyze infrared beam

breaks and quantify locomotor parameters. Locomotor metrics are reported as an average of each of the two light on or two light off periods.

Fiber photometry surgery and calcium imaging

SST-IRES-Cre mice received unilateral infusion (400 nl; A/P: +1.9, D/V: -1.6, M/L: + or - 0.2) of AAV1-hSyn-FLEX-GCaMP6f-WPRE (Addgene # 100833). Following viral infusion, imaging fibers fabricated from 400 μm core 0.48 numerical aperture optic fiber fixed in a 2.5 mm diameter metal ferrule (Doric Lenses, 2.0 mm length) were chronically implanted above the injection site in prelimbic cortex (A/P: +1.9, D/V: -1.5, M/L: + or - 0.2). Fibers were fixed to the skull using luting (Metabond) and dental cement. Mice were returned to their home cage for 4 weeks prior to the start of imaging, which was determined to be the amount of time required to achieve maximal GCaMP6f expression in prefrontal SST-INs.

Mice were habituated to handling (10 min) and patch cord tethering (15 min) for 3 consecutive days prior to the start of imaging experiments. Additionally, calcium signals were streamed during handling and habituation sessions to ensure robust and reproducible signals prior to the start of experiments. Mice were connected to an imaging patch cord (400 μm core; 0.48 numerical aperture, Doric Lenses), which was connected to a 6-port fluorescence minicube (Doric lenses). Blue (465 nm for GCaMP6f excitation, Doric lenses) and violet (405 nm for control artifact fluorescence, Doric Lenses) light was transmitted into the brain at 20–80 μW and kept constant across experimental sessions. Emitted light was passed through a dichroic mirror and a 500–540 nm filter before detection by a visible femtowatt photoreceiver (model 2151, Newport). Analog signals were recorded using a RZ5 processor and a PC equipped with Synapse software (Tucker-Davis Technologies). All conditioning was performed in MedAssociates operant chambers. For paired fear conditioning experiments, mice were tethered to a patch cord and exposed to 4 CS tones (2kHz, 80 dB, 20 s) during two test sessions at 24 hours pre- and post-conditioning. Paired conditioning consisted of 6 pairings of CS with a co-terminating US (0.7 mA, 2 s).

For unpaired conditioning, mice were tethered to the patch cord for both CS and US presentations but were untethered and returned to the home cage for 15 minutes between sessions. After 24 and 48 hours, mice were tethered to the patch cord and exposed to context B and context A, respectively. Calcium signals were sampled at 6kHz and were continuously recorded throughout all tests, including conditioning, CS, and context memory retrieval experiments. All imaging sessions started with a 2 minute baseline period during which calcium signals were recorded prior to the start of the experiment. The start and end of each experimental session as well as CS and US onset and offset timestamps were generated using TTL signals triggered by MedAssociates software to enable precise temporal analysis of calcium signals.

Fiber photometry data analysis

Extraction and analysis of fiber photometry signals was performed using custom code in MATLAB R2018b (original code available on Tucker-Davis Technologies website), following conventional practices reported elsewhere^{53–55}. Demodulated 465 nm and 405 nm signals were digitally filtered and scaled. To account for artifacts resulting from movement,

photobleaching, and autofluorescence, which are reflected by fluctuations in the control 405 nm trace, the control signal was subtracted from the GCaMP6f signal (465 nm). The resulting fitted trace was then normalized to the 405 nm trace ($F/F = (465 \text{ nm signal} - \text{fitted } 405 \text{ nm signal})/\text{fitted } 405 \text{ nm signal}$) and used to calculate changes in fluorescence during behavior. Changes in peak fluorescence in response to a 20-s CS or 2-s US presentation were calculated by normalization to a baseline period of equal duration occurring immediately prior to CS or US onset, respectively. For analysis of inter-CS freezing and contextual freezing epochs, behavior videos were scored to extract freezing onset and offset. Calcium traces associated with the freezing epochs were then aligned and averaged. Changes in fluorescence during the first 2 seconds of freezing, which was the minimum duration of freezing bouts, were compared to a 2 second pre-freezing baseline period immediately preceding the freezing epoch. Percent change in signal ($\% F/F$) was calculated by subtracting the peak fluorescence signal in the test (CS, US, or freezing epoch) period from the peak fluorescence signal in the respective baseline period and dividing the resulting signal by the peak fluorescence signal in the baseline period ($\% F/F = (F_{\text{peakTEST}} - F_{\text{peakBL}})/F_{\text{peakBL}}$). For frequency analysis of calcium-related events, the median absolute deviation of the fitted signal trace was calculated and all transients that exceeded 2.91 deviations were manually counted.

Fear conditioning for electrophysiology

Cued auditory fear conditioning entailed 6 pairings of an auditory tone (CS; 2kHz, 80 dB, 20 s) with a co-terminating footshock (US; 1 mA, 2 s). Control and experimental subjects were selected from the same litter and recordings from these groups were interleaved. All mice were randomly assigned to behavior groups. Naive controls received only cage experience, while unpaired controls were subjected to 6 CS presentations followed 15 min later by 6 US presentations. Mice were sacrificed for electrophysiological analyses at 24 hours after conditioning.

Slice electrophysiology

Mice were deeply anesthetized using isoflurane inhalation prior to decapitation. Acute brain slices were prepared from medial prefrontal cortex at 350 μm thickness on a VT1200S vibratome (Leica Microsystems, Buffalo Grove, IL, USA) in a low sodium sucrose solution bubbled with carbogen (95% O_2 , 5% CO_2) and consisting of (in mM): 210 sucrose, 26.2 NaHCO_3 , 11 glucose, 2.5 KCl, 1 NaH_2PO_4 , 0.5 ascorbate, 4 MgCl_2 , and 0.5 CaCl_2 , and chilled to -3 – -4°C . Slices were transferred to a recovery chamber continuously bubbled with carbogen and containing normal artificial cerebrospinal fluid (ACSF) consisting of (in mM): 119 NaCl, 26.2 NaHCO_3 , 11 glucose, 2.5 KCl, 1 NaH_2PO_4 , 2 MgCl_2 , and 2 CaCl_2 , and warmed to 34°C for 45 min. Following recovery, slices were maintained at room temperature until initiating recordings. Whole-cell electrodes were pulled from borosilicate glass and filled with a low chloride solution (for voltage clamp recordings) consisting of (in mM): 120 Cs-methanesulfonate, 10 HEPES, 10 Na-phosphocreatine, 8 NaCl, 1 QX-314, 0.5 EGTA, 4 Mg-ATP, and 0.4 Na-GTP or a K-based solution (for current clamp recordings) consisting of (in mM): 127.5 K-methanesulfonate, 10 HEPES, 5 KCl, 5 Na-phosphocreatine, 2 MgCl_2 , 0.6 EGTA, 2 Mg-ATP, and 0.3 Na-GTP. Internal solutions were adjusted to pH 7.25 and 290–300 mOsm. Slices were visualized on an upright differential interference

contrast microscope and LED-coupled (Prizmatix, Givat-Shmuel, Israel) 40X objectives were used for the identification of fluorescently-tagged cells as well as optogenetic stimulation. SST- and PV-INs were identified based on tdTomato or EYFP fluorescence, with high membrane resistance ($>100\text{ M}\Omega$) or low capacitance ($<90\text{ pF}$) as confirmation of interneuron identity. Principal excitatory projection neurons (PNs) were identified based on morphology (large pyramidal soma and prominent apical dendrite), with low membrane resistance ($<75\text{ M}\Omega$) or high capacitance ($>100\text{ pF}$) as additional criteria.

Spontaneous excitatory (EPSCs) and inhibitory (IPSCs) postsynaptic currents were isolated by clamping neurons at -60 or 0 mV , respectively, in low-chloride internal solution. A total trace duration of at least 5 minutes was sampled at each potential for spontaneous currents. For paired-pulse measurements, EPSCs were evoked with a bipolar stimulating electrode placed in layer 2 of prelimbic cortex adjacent to the targeted postsynaptic neuron. To stimulate light-evoked transmission, we used TTL-pulsed microscope objective-coupled light-emitting diodes (LEDs, 460 nm , 20 mW/mm^2 , Prizmatix). This intensity evoked maximal response amplitudes at a pulse duration of 1 ms . Spontaneous postsynaptic currents (Fig. 1 and Supplementary Fig. 1), paired-pulse analysis (Fig. 1) and light-evoked compound currents (Supplementary Fig. 13) were conducted in standard ACSF. For more stringent isolation of monosynaptic currents during optic stimulation (Fig. 5; Fig. 7), recordings were conducted in the presence of $1\text{ }\mu\text{M}$ tetrodotoxin (Abcam) and $100\text{ }\mu\text{M}$ 4-aminopyrimidine (Abcam), which results in complete elimination of polysynaptic activity^{12,13,30}.

Data were low-pass filtered at 3 kHz (evoked) or 10 kHz (spontaneous) and acquired at 10 kHz using Multiclamp 700B (Molecular Devices, San Jose, CA, USA) and pClamp 10 software (Molecular Devices). Mice were randomly assigned to behavior groups and the experimenter was blind to cell type and experimental condition during analysis of evoked (Clampfit 10, Molecular Devices) and spontaneous currents (MiniAnalysis, Synaptosoft, Fort Lee, NJ, USA). Neurons were excluded prior to analysis if they did not meet cell-type criteria for passive membrane properties (as detailed above) or yielded unstable current or voltage traces. These rejected neurons accounted for a very small number of recordings.

cFos immunofluorescence

For analysis of CS-evoked cFos (Supplementary Fig. 5–6; Fig. 8i–j; Supplementary Fig. 17), mice were subjected to 4 CS presentations in context B. For analysis of cFos induction by SST-IN photoexcitation (Fig. 8a–h), ChR2 or eYFP-expressing mice were subjected to 6 bouts of 20 s stimulation (473 nm , 10 ms pulse, 20 Hz). Ninety minutes following CS or light stimuli, mice were deeply anesthetized via isoflurane inhalation and transcardially perfused with phosphate buffered saline (PBS) followed by 4% paraformaldehyde (PFA) in PBS (pH 7.45). Brains were post-fixed in PFA for 14–16 hours post perfusion and sectioned in $50\text{ }\mu\text{m}$ thick slices on the coronal plane on a VT1000S vibratome (Leica).

Immunofluorescence staining against cFos was conducted on floating sections using a rabbit anti-cFos primary antibody (1:1000, Millipore, ABE457)⁵⁶. Fluorescence-conjugated secondary antibodies included goat anti-rabbit conjugated to FITC (1:500, 111-095-003) and goat anti-rabbit conjugated to Alexa 647 (1:500, 111-605-003) and were purchased from Jackson ImmunoResearch (West Grove, PA, USA). Slices were blocked in 2% goat serum in

0.3% Tween-20 PBS for 1 hour at room temperature. The primary antibody was incubated overnight at 4°C in 2% goat serum in 0.3% Tween-20 PBS. Following PBS washes, slices were incubated with the secondary antibody in 2% goat serum in 0.3% Tween-20 for 2 hours at room temperature. Following additional PBS washes, slices were incubated for 7 minutes at room temperature with filtered 1 mg/mL DAPI solution, followed by thorough PBS washes. Slices were then mounted with Prolong Antifade Gold mountant medium (Life Technologies, Grand Island, NY, USA) and imaged on a Zeiss confocal microscope operating Zeiss Zen software (Carl Zeiss Microscopy, Jena, Germany). tdTomato+ and eYFP + SST-INs, as well as cFos+ nuclei, were manually quantified using the Cell Counter plugin in ImageJ (National Institutes of Health, Bethesda, MD, USA) while blind to experimental condition.

Statistical analysis

Prior to parametric statistical analysis, the Shapiro-Wilk test and Levene's test were used to establish normality of data and homogeneity of variance, respectively. Failing these assumptions or in cases in which group sizes might be too small to establish normality, we utilized non-parametric statistical comparisons. All figures utilizing bar graphs contain individual sample data, means, and standard error bars. Parameters used for box plots are included in corresponding figure legends. For reported effects, statistical power exceeded a minimum of 0.8, and in most cases 0.9. P-values obtained from network analysis of cFos+ cells were corrected for a false discovery rate of 10% using the method of Benjamini and Hochberg. Power analyses were conducted at the outset of the project to estimate sample sizes for electrophysiological recordings and freezing behavior based on previous experiments published by our laboratory. Statistical analysis and graphing were conducted in Graphpad Prism 6 (San Diego, CA) and OriginPRO 2016 (OriginLab, Northampton, MA).

All additional information regarding experimental subjects, statistical parameters, software and code, study design, and materials and methods can be found in the associated Life Sciences Reporting Summary document.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Bayshtok and E. Beckett for expert technical assistance, behavioral scoring and assistance with experimenter blinding; C. Román-Ortiz for help with optogenetic behavioral manipulations; and D. Cai for comments on the manuscript. These experiments were supported by funds from National Institute of Mental Health (NIMH) grants RO1 MH105414, RO1 MH116145, and R21 MH114170 to R.L.C, in addition to F32 MH115688 to K.A.C.

References

1. Holtmaat A & Caroni P Functional and structural underpinnings of neuronal assembly formation in learning. *Nat Neurosci* 19, 1553–1562, doi:10.1038/nn.4418 (2016). [PubMed: 27749830]
2. Kandel ER, Dudai Y & Mayford MR The molecular and systems biology of memory. *Cell* 157, 163–186, doi:10.1016/j.cell.2014.03.001 (2014). [PubMed: 24679534]

3. Josselyn SA, Kohler S & Frankland PW Finding the engram. *Nat Rev Neurosci* 16, 521–534, doi: 10.1038/nrn4000 (2015). [PubMed: 26289572]
4. Courtin J et al. Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* 505, 92–96, doi:10.1038/nature12755 (2014). [PubMed: 24256726]
5. Rashid AJ et al. Competition between engrams influences fear memory formation and recall. *Science* 353, 383–387, doi:10.1126/science.aaf0594 (2016). [PubMed: 27463673]
6. Siwani S et al. OLMalpha2 Cells Bidirectionally Modulate Learning. *Neuron* 99, 404–412 e403, doi:10.1016/j.neuron.2018.06.022 (2018). [PubMed: 29983324]
7. Stefanelli T, Bertollini C, Luscher C, Muller D & Mendez P Hippocampal Somatostatin Interneurons Control the Size of Neuronal Memory Ensembles. *Neuron* 89, 1074–1085, doi: 10.1016/j.neuron.2016.01.024 (2016). [PubMed: 26875623]
8. Wolff SB et al. Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* 509, 453–458, doi:10.1038/nature13258 (2014). [PubMed: 24814341]
9. Corcoran KA & Quirk GJ Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *J Neurosci* 27, 840–844, doi:10.1523/JNEUROSCI.5327-06.2007 (2007). [PubMed: 17251424]
10. Rudy B, Fishell G, Lee S & Hjerling-Leffler J Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev Neurobiol* 71, 45–61, doi:10.1002/dneu.20853 (2011). [PubMed: 21154909]
11. Kim D et al. Distinct Roles of Parvalbumin- and Somatostatin-Expressing Interneurons in Working Memory. *Neuron* 92, 902–915, doi:10.1016/j.neuron.2016.09.023 (2016). [PubMed: 27746132]
12. Arruda-Carvalho M & Clem RL Pathway-selective adjustment of prefrontal-amygdala transmission during fear encoding. *J Neurosci* 34, 15601–15609, doi:10.1523/JNEUROSCI.2664-14.2014 (2014). [PubMed: 25411488]
13. Lucas EK, Jegarl AM, Morishita H & Clem RL Multimodal and Site-Specific Plasticity of Amygdala Parvalbumin Interneurons after Fear Learning. *Neuron* 91, 629–643, doi:10.1016/j.neuron.2016.06.032 (2016). [PubMed: 27427462]
14. Graziane N & Dong Y Electrophysiological analysis of synaptic transmission. (Humana Press : Springer, 2016).
15. Herry C & Johansen JP Encoding of fear learning and memory in distributed neuronal circuits. *Nat Neurosci* 17, 1644–1654, doi:10.1038/nn.3869 (2014). [PubMed: 25413091]
16. Amano T, Unal CT & Pare D Synaptic correlates of fear extinction in the amygdala. *Nat Neurosci* 13, 489–494, doi:10.1038/nn.2499 (2010). [PubMed: 20208529]
17. Asede D, Bosch D, Luthi A, Ferraguti F & Ehrlich I Sensory inputs to intercalated cells provide fear-learning modulated inhibition to the basolateral amygdala. *Neuron* 86, 541–554, doi:10.1016/j.neuron.2015.03.008 (2015). [PubMed: 25843406]
18. Namburi P et al. A circuit mechanism for differentiating positive and negative associations. *Nature* 520, 675–678, doi:10.1038/nature14366 (2015). [PubMed: 25925480]
19. Pattwell SS, Bath KG, Casey BJ, Ninan I & Lee FS Selective early-acquired fear memories undergo temporary suppression during adolescence. *Proc Natl Acad Sci U S A* 108, 1182–1187, doi:10.1073/pnas.1012975108 (2011). [PubMed: 21220344]
20. Tsvetkov E, Carlezon WA, Benes FM, Kandel ER & Bolshakov VY Fear conditioning occludes LTP-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. *Neuron* 34, 289–300 (2002). [PubMed: 11970870]
21. Zhou Y et al. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. *Nat Neurosci* 12, 1438–1443, doi:10.1038/nn.2405 (2009). [PubMed: 19783993]
22. Clem RL & Huganir RL Calcium-permeable AMPA receptor dynamics mediate fear memory erasure. *Science* 330, 1108–1112, doi:10.1126/science.1195298 (2010). [PubMed: 21030604]
23. Pi HJ et al. Cortical interneurons that specialize in disinhibitory control. *Nature* 503, 521–524, doi: 10.1038/nature12676 (2013). [PubMed: 24097352]
24. Fenno LE et al. Targeting cells with single vectors using multiple-feature Boolean logic. *Nat Methods* 11, 763–772, doi:10.1038/nmeth.2996 (2014). [PubMed: 24908100]

25. Senn V et al. Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* 81, 428–437, doi:10.1016/j.neuron.2013.11.006 (2014). [PubMed: 24462103]
26. Herry C et al. Switching on and off fear by distinct neuronal circuits. *Nature* 454, 600–606, doi: 10.1038/nature07166 (2008). [PubMed: 18615015]
27. Burgos-Robles A et al. Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nat Neurosci* 20, 824–835, doi:10.1038/nn.4553 (2017). [PubMed: 28436980]
28. Klavir O, Prigge M, Sarel A, Paz R & Yizhar O Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. *Nat Neurosci* 20, 836–844, doi:10.1038/nn.4523 (2017). [PubMed: 28288126]
29. Little JP & Carter AG Synaptic mechanisms underlying strong reciprocal connectivity between the medial prefrontal cortex and basolateral amygdala. *J Neurosci* 33, 15333–15342, doi:10.1523/JNEUROSCI.2385-13.2013 (2013). [PubMed: 24068800]
30. Cruikshank SJ, Urabe H, Nurmikko AV & Connors BW Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. *Neuron* 65, 230–245, doi:10.1016/j.neuron.2009.12.025 (2010). [PubMed: 20152129]
31. McGarry LM & Carter AG Inhibitory Gating of Basolateral Amygdala Inputs to the Prefrontal Cortex. *J Neurosci* 36, 9391–9406, doi:10.1523/JNEUROSCI.0874-16.2016 (2016). [PubMed: 27605614]
32. Lee S, Kruglikov I, Huang ZJ, Fishell G & Rudy B A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat Neurosci* 16, 1662–1670, doi:10.1038/nn.3544 (2013). [PubMed: 24097044]
33. Garcia-Junco-Clemente P et al. An inhibitory pull-push circuit in frontal cortex. *Nat Neurosci* 20, 389–392, doi:10.1038/nn.4483 (2017). [PubMed: 28114295]
34. Ibrahim LA et al. Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. *Neuron* 89, 1031–1045, doi:10.1016/j.neuron.2016.01.027 (2016). [PubMed: 26898778]
35. Letzkus JJ et al. A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–335, doi:10.1038/nature10674 (2011). [PubMed: 22158104]
36. Beas BS et al. The locus coeruleus drives disinhibition in the midline thalamus via a dopaminergic mechanism. *Nat Neurosci* 21, 963–973, doi:10.1038/s41593-018-0167-4 (2018). [PubMed: 29915192]
37. Do-Monte FH, Quinones-Laracuenta K & Quirk GJ A temporal shift in the circuits mediating retrieval of fear memory. *Nature* 519, 460–463, doi:10.1038/nature14030 (2015). [PubMed: 25600268]
38. Stamatakis AM & Stuber GD Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nat Neurosci* 15, 1105–1107, doi:10.1038/nn.3145 (2012). [PubMed: 22729176]
39. Tovote P et al. Midbrain circuits for defensive behaviour. *Nature* 534, 206–212, doi:10.1038/nature17996 (2016). [PubMed: 27279213]
40. Johnson PL & Shekhar A Panic-prone state induced in rats with GABA dysfunction in the dorsomedial hypothalamus is mediated by NMDA receptors. *J Neurosci* 26, 7093–7104, doi: 10.1523/JNEUROSCI.0408-06.2006 (2006). [PubMed: 16807338]
41. Arruda-Carvalho M, Wu WC, Cummings KA & Clem RL Optogenetic Examination of Prefrontal-Amygdala Synaptic Development. *J Neurosci* 37, 2976–2985, doi:10.1523/JNEUROSCI.3097-16.2017 (2017). [PubMed: 28193691]
42. Xu W & Sudhof TC A neural circuit for memory specificity and generalization. *Science* 339, 1290–1295, doi:10.1126/science.1229534 (2013). [PubMed: 23493706]
43. Yavorska I & Wehr M Somatostatin-Expressing Inhibitory Interneurons in Cortical Circuits. *Front Neural Circuits* 10, 76, doi:10.3389/fncir.2016.00076 (2016). [PubMed: 27746722]
44. Munoz W, Tremblay R, Levenstein D & Rudy B Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science* 355, 954–959, doi:10.1126/science.aag2599 (2017). [PubMed: 28254942]

45. Li K, Nakajima M, Ibanez-Tallon I & Heintz N A Cortical Circuit for Sexually Dimorphic Oxytocin-Dependent Anxiety Behaviors. *Cell* 167, 60–72 e11, doi:10.1016/j.cell.2016.08.067 (2016). [PubMed: 27641503]
46. Nakajima M, Gorlich A & Heintz N Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. *Cell* 159, 295–305, doi:10.1016/j.cell.2014.09.020 (2014). [PubMed: 25303526]
47. Kluge C, Stoppel C, Szinyei C, Stork O & Pape HC Role of the somatostatin system in contextual fear memory and hippocampal synaptic plasticity. *Learn Mem* 15, 252–260, doi:10.1101/lm.793008 (2008). [PubMed: 18391186]
48. Einstein EB et al. Somatostatin signaling in neuronal cilia is critical for object recognition memory. *J Neurosci* 30, 4306–4314, doi:10.1523/JNEUROSCI.5295-09.2010 (2010). [PubMed: 20335466]
49. Xu H et al. A Disinhibitory Microcircuit Mediates Conditioned Social Fear in the Prefrontal Cortex. *Neuron* 102, 668–682 e665, doi:10.1016/j.neuron.2019.02.026 (2019). [PubMed: 30898376]
50. Keum S et al., A Missense Variant at the *Nrxn3* Locus Enhances Empathy Fear in the Mouse. *Neuron* 98, 588 (5 2, 2018). [PubMed: 29681532]
51. Yaeger CE, Ringach DL, Trachtenberg JT, Neuromodulatory control of localized dendritic spiking in critical period cortex. *Nature* 567, 100 (3, 2019). [PubMed: 30787434]
52. Kvitsiani D et al., Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* 498, 363 (6 20, 2013). [PubMed: 23708967]
53. Calipari ES et al., In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc Natl Acad Sci U S A* 113, 2726 (3 8, 2016). [PubMed: 26831103]
54. Gunaydin LA et al., Natural neural projection dynamics underlying social behavior. *Cell* 157, 1535 (6 19, 2014). [PubMed: 24949967]
55. Beas BS et al., The locus coeruleus drives disinhibition in the midline thalamus via a dopaminergic mechanism. *Nat Neurosci* 21, 963 (7, 2018). [PubMed: 29915192]
56. Lagerlöf O et al., The nutrient sensor OGT in PVN neurons regulates feeding. *Science* 351, 6279 (3 2016).

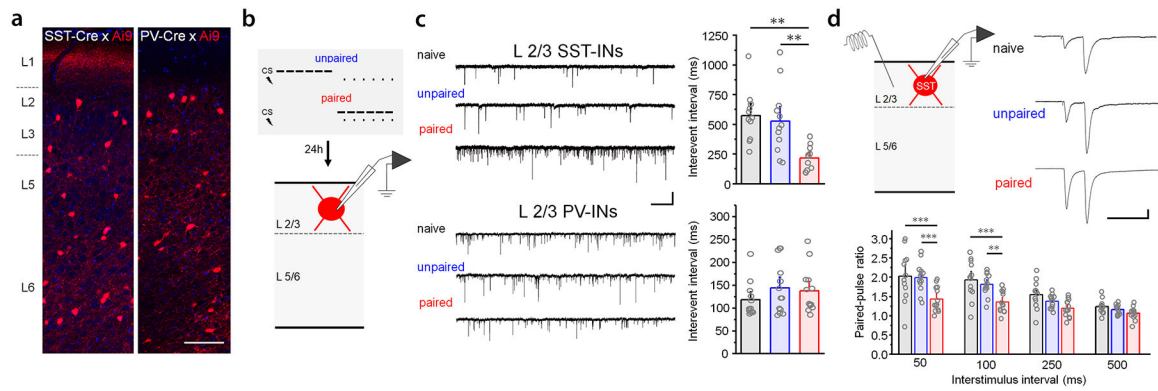


Fig. 1. Potentiation of synaptic transmission in prefrontal SST-INs after cued fear learning.
a, Prelimbic (PL) SST-INs and PV-INs were identified by tdTomato expression in SST-IRES-Cre/ Ai9 or PV-IRES-Cre/ Ai9 mice. Scale = 200 μ m. **b**, Subjects were either cage-experienced (naïve), or trained using paired or unpaired presentations of CS (2 KHz, 80 dB, 20 s) and US (1 mA, 2-s), followed 24 hrs later by whole-cell recording in acute brain slices. **c**, Recordings of sEPSCs were obtained from L2/3 SST-INs (upper) and PV-INs (lower). Example raw traces and mean inter-event intervals are depicted for experimental and control groups. SST-INs: $F_{2,31} = 9.03$, $p = 8.11 \times 10^{-4}$, 1-way ANOVA; naïve, $n = 11$ cells (7 slices from 3 mice); unpaired $n = 12$ cells (9 slices from 3 mice); paired $n = 11$ cells (12 slices from 4 mice). PV-INs: $F_{2,45} = 0.52$, $p = 0.60$, 1-way ANOVA; naïve, $n = 13$ cells (9 slices from 3 mice); unpaired, $n = 13$ cells (8 slices from 3 mice); paired, $n = 13$ cells (11 slices from 4 mice). Scale = 20 pA \times 1s. **d**, EPSC recordings in L2/3 SST-INs during paired-pulse stimulation. Paired pulse ratio: $F_{6,66} = 3.46$, $p = 0.0049$, interaction between ratio and training, 2-way repeated measures ANOVA; naïve, $n = 13$ cells (7 slices from 3 mice); unpaired $n = 13$ cells (6 slices from 3 mice); paired, $n = 13$ cells (6 slices from 3 mice). Scale = 100 pA \times 100 ms. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ Tukey's post-hoc test. Bar graphs depict mean \pm SE.

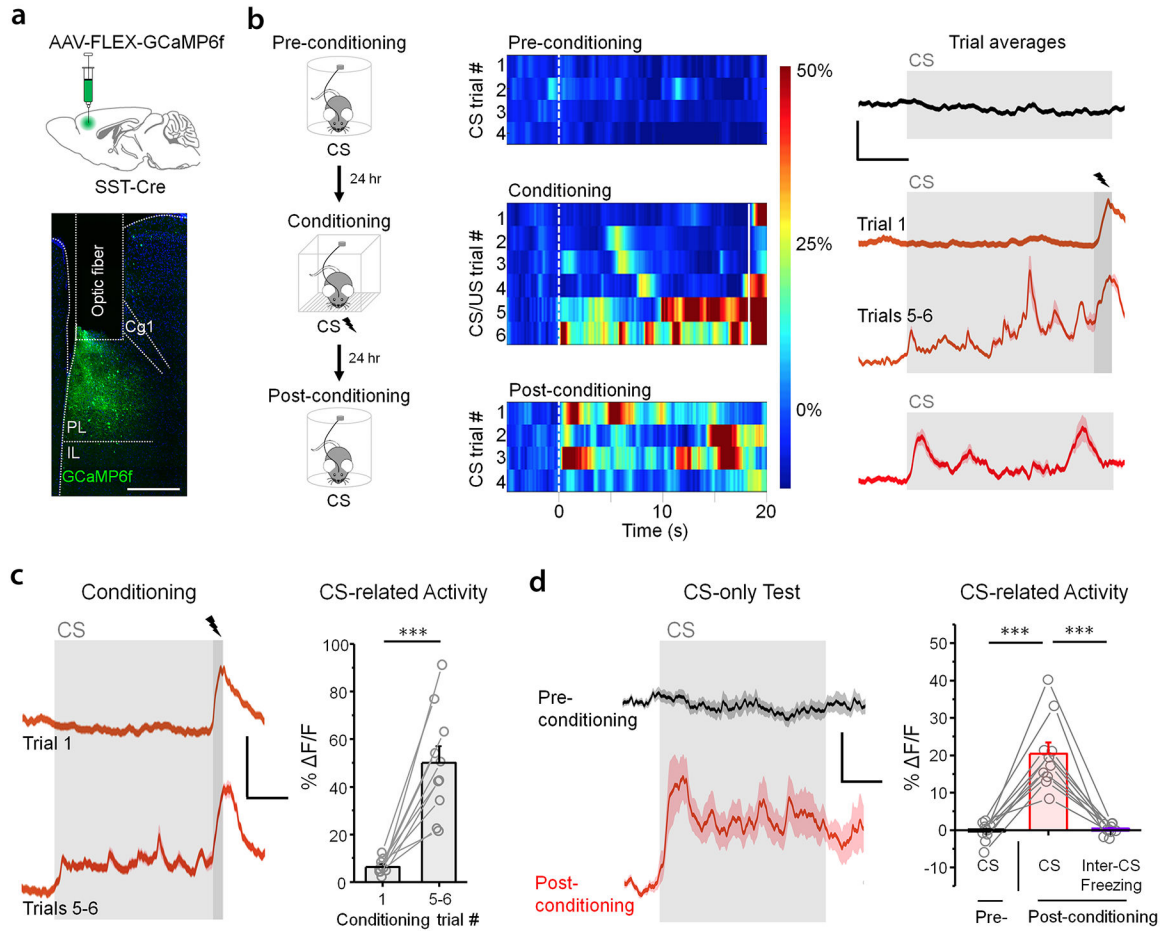


Fig. 2. **CS-related SST-IN activity increases in tandem with memory acquisition.** **a**, For Ca^{2+} -based imaging of SST-IN activity, SST-IRES-Cre mice ($n = 10$) received injections of conditional vector encoding GCaMP6f and were implanted with a single optic fiber ($400\ \mu m$ core diameter) directed at PL. Scale = $500\ \mu m$. **b**, After surgical recovery, CS-related freezing and Ca^{2+} -dependent fluorescence signal were monitored before, during and after CS-US pairing, which entailed 6 co-terminating trials of CS (2 KHz, 20 s, 80 dB) and US (0.7 mA foot shock, 2 s). Pre- and post-conditioning CS tests consisted of 4 CS trials in a context distinct from the training arena. Heat maps (middle) and mean Ca^{2+} traces (right) \pm SE (shaded area) depict SST-IN responses to all CS and US trials for a representative animal. Scale = $30\% \ F/F \times 5\ s$. **c**, Group averaged Ca^{2+} traces (left) \pm SE (shaded area) for conditioning CS/US trials and mean change in CS-associated peak fluorescence ($\% \ F/F$) for trials 1 and trials 5–6. Effect of trial: $n = 10$ mice, $t_9 = 5.95$, $p = 2.15 \times 10^{-4}$, two-sided paired t-test. Scale = $50\% \ F/F \times 5\ s$. **d**, Group averaged Ca^{2+} traces (left) \pm standard error (shaded area). Right, comparison of percent F/F between pre- and post-conditioning CS presentations, as well as freezing-related epochs that occurred independent of the CS during the inter-trial intervals of the post-conditioning test. CS-related activity: $n = 10$ mice, $F_{2,18} = 42.82$, $p = 1.44 \times 10^{-7}$, 1-way repeated measures ANOVA. Experiment was performed in 3 independent cohorts and mice were pooled together for analysis. Traces represent the

average of 4 CS trials. Scale = 10% F/F \times 5 s. *** $p < 0.001$ by two-sided paired t-test (c) or Tukey's post-hoc test (d). Bar graphs depict mean \pm SE.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

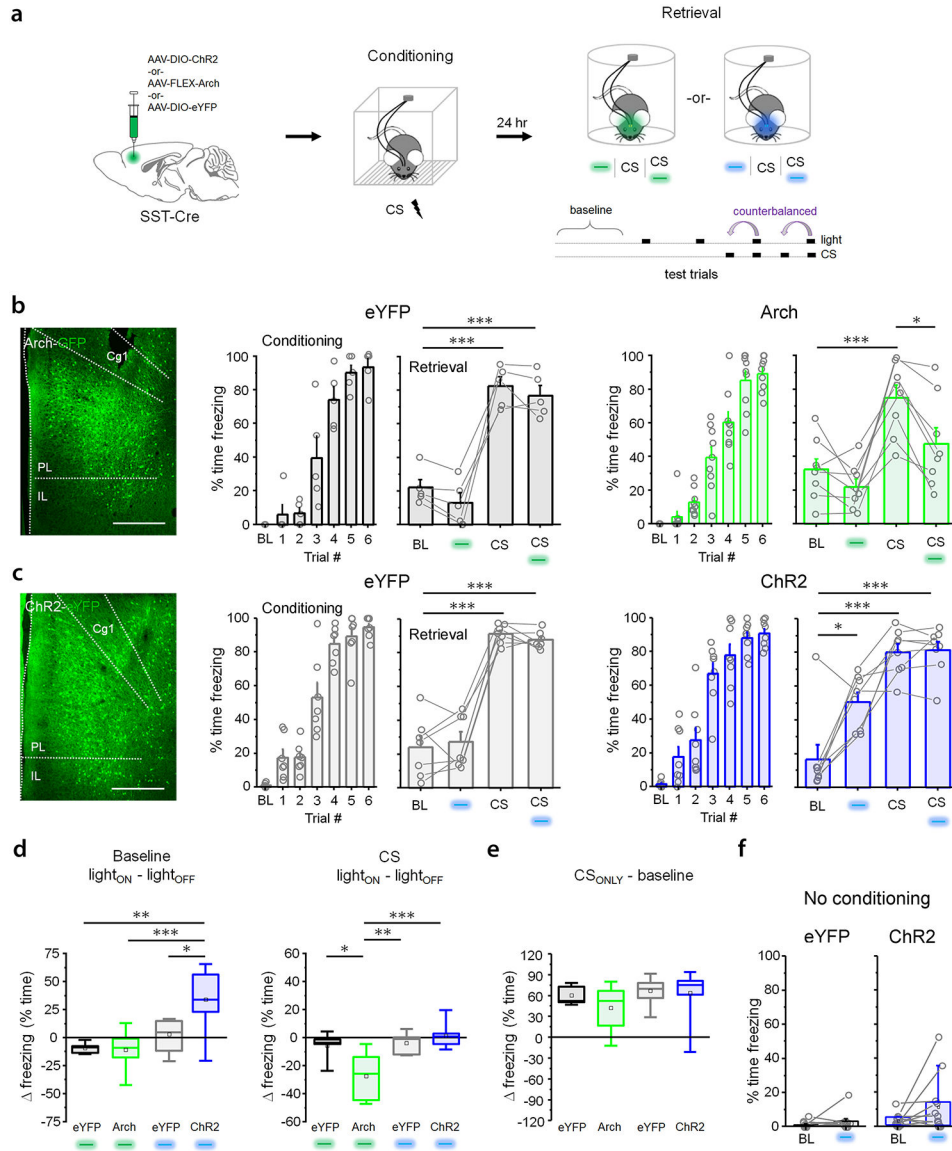


Fig. 3. SST-IN activity is necessary for memory expression and induces freezing in conditioned mice.

a. For *in vivo* activity manipulation, SST-IRES-Cre mice received injections of conditional Arch, ChR2, or eYFP control vectors and were implanted with optic ferrules directed at PL. After surgical recovery, all subjects underwent auditory fear conditioning in the absence of optic illumination. Freezing was quantified 24 hrs later during independent and simultaneous presentation of CS and photostimulation in a distinct context. **b.** Modulation of freezing by CS and light (532 nm, constant, 20 s epochs) in Arch (green) or eYFP control mice (black). Arch: $F_{3,21} = 11.3$, $p = 1.29 \times 10^{-4}$, 1-way repeated measures ANOVA; $n = 8$ mice. eYFP: $F_{3,12} = 68.6$, $p = 8.03 \times 10^{-8}$, repeated measures ANOVA; $n = 5$ mice. Scale = 500 μm . Experiment was performed in three independent cohorts and mice were pooled together for analysis. **c.** Modulation of freezing by CS and light (473 nm, 10 ms pulses at 20 Hz) in ChR2 (blue) or eYFP control mice (gray). ChR2: $F_{3,21} = 19.3$, $p = 3.06 \times 10^{-6}$, 1-way repeated measures ANOVA; $n = 9$ mice. eYFP: $F_{3,18} = 87.7$, $p = 6.20 \times 10^{-11}$, 1-way

repeated measures ANOVA; $n = 7$ mice. Scale = 500 μm . Experiment was performed in four independent cohorts and mice were pooled together for analysis. **d**, Mean change in freezing induced by photostimulation during baseline and CS periods, as compared across all light/opsin combinations used in **(b)** and **(c)**. Photostimulation effect on baseline freezing ($\text{light}_{\text{ON}} - \text{light}_{\text{OFF}}$): $F_{3,24} = 8.88$, $p = 3.87 \times 10^{-4}$, 1-way ANOVA. Photostimulation effect on CS-evoked freezing ($\text{light}_{\text{ON}} - \text{light}_{\text{OFF}}$): $F_{3,24} = 9.61$, $p = 2.37 \times 10^{-4}$, 1-way ANOVA. **e**, Mean change in freezing induced by CS presentation, as compared across all experimental groups used in **(b)** and **(c)**. CS-evoked change in freezing ($\text{CS}_{\text{ONLY}} - \text{baseline}$): $F_{3,24} = 1.12$, $p = 0.36$, 1-way ANOVA. **f**, Freezing behavior during light-only trials (473 nm, 10 ms pulses, 20 Hz) in naïve (unconditioned) mice expressing either ChR2 or eYFP control vector. ChR2: $W = 17$, $p = 0.088$, two-sided Wilcoxon signed ranks test; $n = 12$ mice. eYFP: $W = 12$, $p = 0.46$, two-sided Wilcoxon signed ranks test; $n = 11$ mice. Experiment was performed in four independent cohorts and mice were pooled together for analysis. BL = baseline period. CS = conditioned stimulus (2 kHz tone, 80 db, 20 s duration). Cg1 = cingulate area 1. IL = infralimbic cortex. Blue/ green boxes = laser illumination. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ Tukey's post-hoc test (**b**, **c**, **d**). Bar graphs depict mean \pm SE. Box plots depict median (center line), mean (black box), quartiles, and 10–90% range (whiskers).

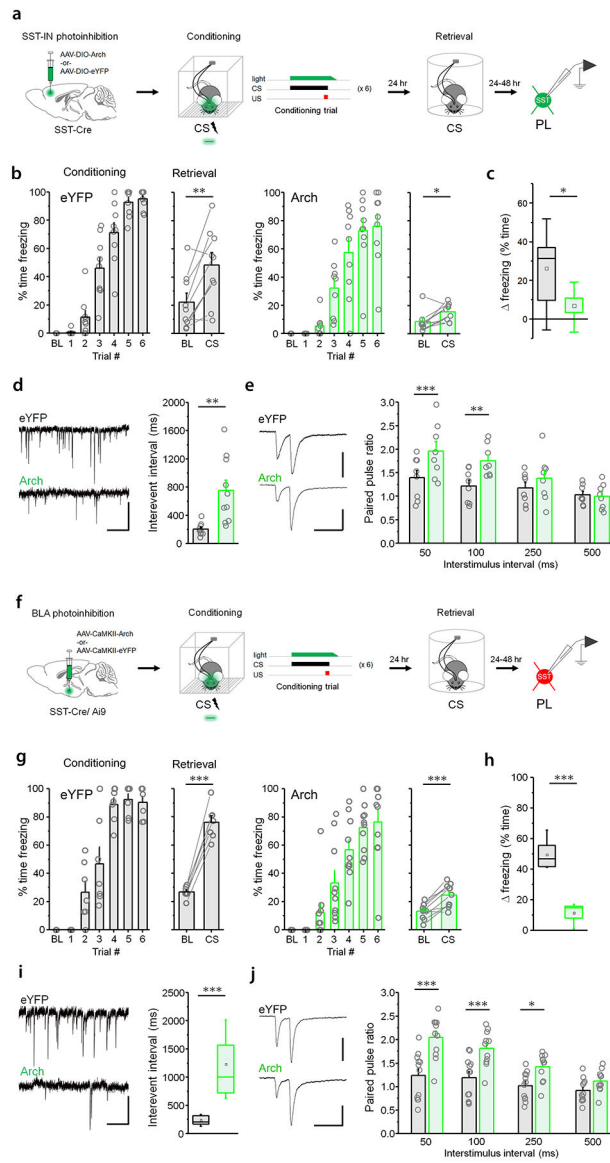


Fig. 4. SST-IN activation and plasticity mediates memory formation.

a, For *in vivo* manipulation of prelimbic SST-INS, SST-IRES-Cre mice received injections of conditional Arch or eYFP control vectors and were implanted with optic ferrules directed at PL. After surgical recovery, all subjects underwent auditory fear conditioning, during which light stimulation (532 nm, constant, 20 s epoch, ramp offset) coincided with each of 6 CS-US trials. Freezing was quantified 24 hrs later during presentation of 4 CS trials context distinct from the training arena. During the 24–48 hours subsequent to the retrieval test, *ex vivo* recordings were obtained from eYFP-positive neurons (SST-INS) located not more than 1 mm below the tip of the optic fiber in layer 2/3 of acute brain slices. **b**, Modulation of freezing by CS retrieval in Arch (green) or eYFP control mice (black). Arch retrieval: $t_8 = 2.71$, $p = 0.026$, two-sided paired t-test; $n = 9$ mice. eYFP retrieval: $t_8 = 3.78$, $p = 0.0054$, paired t-test; $n = 9$ mice. **c**, Mean change in freezing induced by CS presentation during memory retrieval for experimental groups in (**b**). Arch versus eYFP: $t_{16} = 2.60$, $p = 0.019$,

two-sided unpaired t-test. **d**, Example raw traces and interevent intervals (IEIs) for sEPSCs in Arch versus eYFP mice. IEI: $t_{16} = 2.60$, $p = 0.019$, two-sided unpaired t-test; Arch, $n = 10$ cells (5 slices from 3 mice); eYFP, $n = 10$ cells (5 slices from 3 mice). Scale = $10 \text{ pA} \times 0.5 \text{ s}$. **e**, EPSC recordings in SST-INs during local paired-pulse electrical stimulation. Example traces collected at 50 ms interstimulus interval. Paired pulse ratio: $F_{3,21} = 5.55$, $p = 0.0058$, interaction between ratio and training, 2-way repeated measures ANOVA; Arch, $n = 8$ cells (4 slices from 3 mice); eYFP, $n = 8$ cells (4 slices from 3 mice). Scale = $40 \text{ pA} \times 100 \text{ ms}$. **f**, Experimental design for *in vivo* manipulation of BLA projection neurons was the same as in **(a)**, with the exception that vector injections and light stimulation were delivered to the BLA, and *ex vivo* recordings were obtained from Tomato-positive SST-INs in the prelimbic cortex of SST-IRES-Cre/ Ai9 mice. **g**, Modulation of freezing by CS retrieval in Arch- (green) or eYFP control mice (black). Arch retrieval: $t_9 = 5.69$, $p = 2.98 \times 10^{-4}$, two-sided paired t-test; $n = 10$ mice. eYFP retrieval: $t_6 = 14.74$, 6.14×10^{-6} , paired t-test; $n = 7$ mice. **h**, Mean change in freezing induced by CS presentation during memory retrieval for experimental groups in **(b)**. Arch versus eYFP: $U = 0$, $p = 1.03 \times 10^{-4}$, two-sided Mann-Whitney U test. **i**, Example raw traces and interevent intervals (IEIs) for sEPSCs in Arch versus eYFP mice. IEI: $U = 0$, $p = 9.97 \times 10^{-4}$, two-sided Mann-Whitney U test; Arch, $n = 14$ cells (6 slices from 3 mice); eYFP, $n = 13$ cells (8 slices from 4 mice). Scale = $10 \text{ pA} \times 0.5 \text{ s}$. Scale = $10 \text{ pA} \times 0.5 \text{ s}$. **j**, EPSC recordings in SST-INs during local paired-pulse electrical stimulation. Example traces collected at 50 ms interstimulus interval. Paired pulse ratio: $F_{3,30} = 11.18$, $p = 4.31 \times 10^{-5}$, interaction between ratio and training, 2-way repeated measures ANOVA; Arch, $n = 12$ cells (6 slices from 3 mice); eYFP, $n = 12$ cells (6 slices from 3 mice). Scale = $50 \text{ pA} \times 100 \text{ ms}$. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by two-sided paired t-test (**b**, **g**), two-sided unpaired t-test (**c**, **d**), two-sided Mann-Whitney U test (**h**, **i**), or Tukey's post-hoc test (**j**). Bar graphs depict mean \pm SE. Box plots depict median (center line), mean (black box), quartiles, and 10–90% range (whiskers).

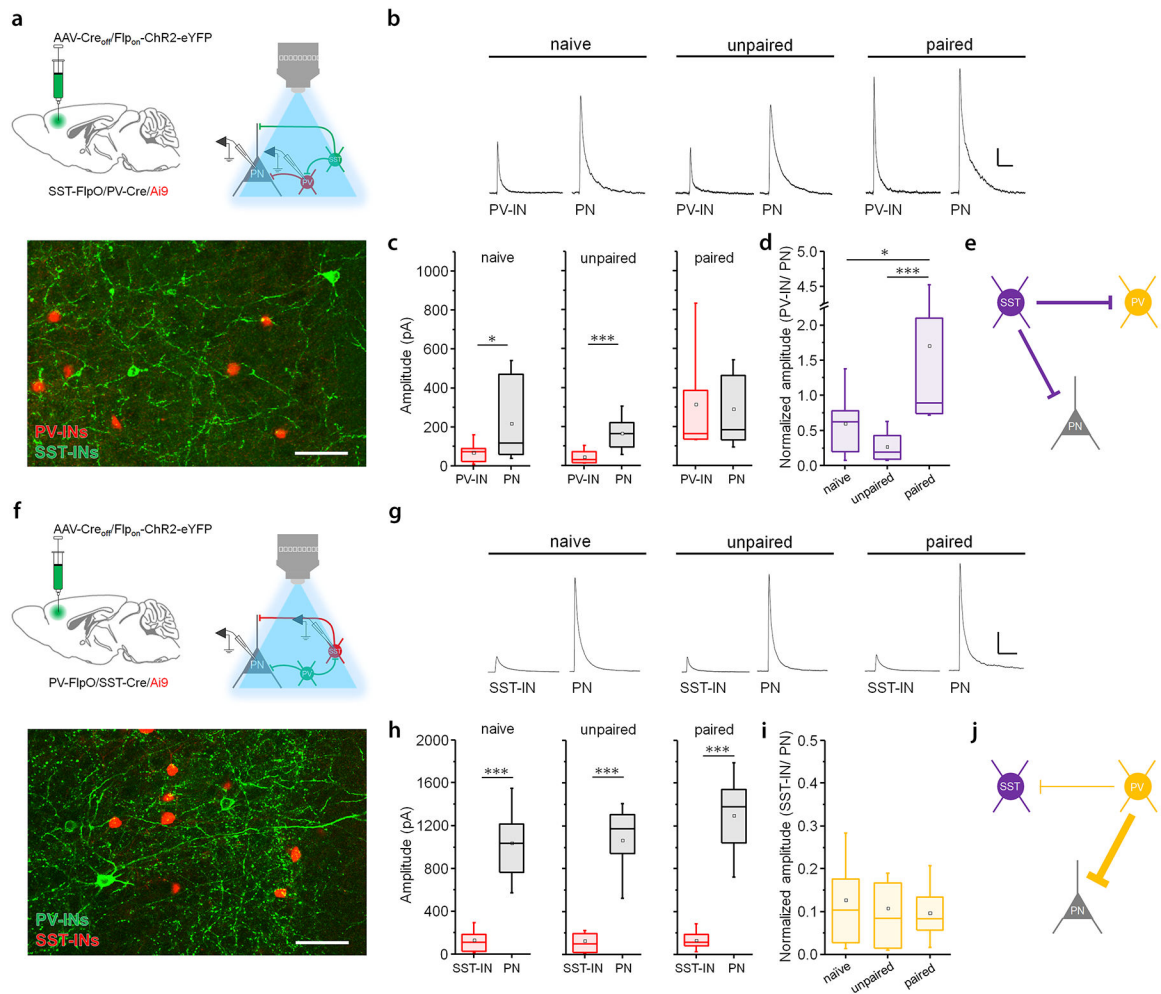


Fig. 5. SST-INs elicit relatively potent inhibition of PV-INs while PV-INs preferentially inhibit PNs.

a, To isolate monosynaptic responses to SST-IN photostimulation in PV-INs, we first infused a Flp-ON, Cre-OFF INTRASECT ChR2 vector into SST-IRES-FlpO/ PV-IRES-Cre/ Ai9 triple transgenic mice. At 24 hours after CS-US pairing, recordings were obtained from Tomato⁺ PV-INs as well as surrounding PNs. Scale = 50 μ m. **b**, Example IPSC traces. Scale = 200 pA \times 1 s. **c**, Amplitude of IPSCs resulting from SST-IN photoexcitation (460 nm, 1 ms pulse, 0.1 Hz) in slices from naïve mice (8 slices from 4 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 46.5$, $p = 0.017$, two-sided Mann-Whitney U test; PV-IN, $n = 14$ cells; PN, $n = 14$ cells. Effect of cell type (unpaired): $U = 17$, $p = 8.22 \times 10^{-4}$, two-sided Mann-Whitney U test; PV-IN, $n = 12$ cells, PN, $n = 12$ cells. Effect of cell type (paired): $U = 108.5$, $p = 0.88$, two-sided Mann-Whitney U test; PV-IN, $n = 15$ cells; PN, $n = 15$ cells. **d**, Amplitude of IPSCs in PV-INs normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 20.63$ (2), $p = 3.31 \times 10^{-5}$, Kruskal-Wallis ANOVA. **e**, Relative strength of SST-IN transmission onto PV-INs versus PNs in conditioned mice. **f**, Interrogation of monosynaptic connections from PV-INs onto SST-INs and PNs. Flp-ON, Cre-OFF INTRASECT ChR2 vector into PV-IRES-FlpO/ SST-IRES-Cre/ Ai9 triple

transgenic mice. **g**, Example IPSC traces. Scale = 200 pA × 1s. **h**, Amplitude of IPSCs resulting from PV-IN photoexcitation (460 nm, 1 ms pulse, 0.1 Hz) in slices from naïve mice (6 slices from 3 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (6 slices from 3 mice). Effect of cell type (naïve): $U = 6$, $p = 1.15 \times 10^{-5}$, two-sided Mann-Whitney U test; SST-IN, $n = 14$ cells; PN, $n = 15$ cells. Effect of cell type (unpaired): $U = 3$, $p = 1.04 \times 10^{-5}$, two-sided Mann-Whitney U test; PV-IN, $n = 12$ cells, PN, $n = 11$ cells. Effect of cell type (paired): $U = 6$, $p = 1.15 \times 10^{-5}$, two-sided Mann-Whitney U test; PV-IN, $n = 12$ cells; PN, $n = 13$ cells. **i**, Amplitude of IPSCs in PV-INs normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 0.33$ (2), $p = 0.85$, Kruskal-Wallis ANOVA. **j**, Relative strength of PV-IN transmission onto SST-INs versus PNs in conditioned mice. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by Mann-Whitney U test (**c**, **h**) or Dunn's post-hoc test (**d**). Box plots depict median (center line), mean (black box), quartiles, and 10–90% range (whiskers). During optic stimulation, 100% of postsynaptic cells that were sampled in these analyses exhibited synaptic responses.

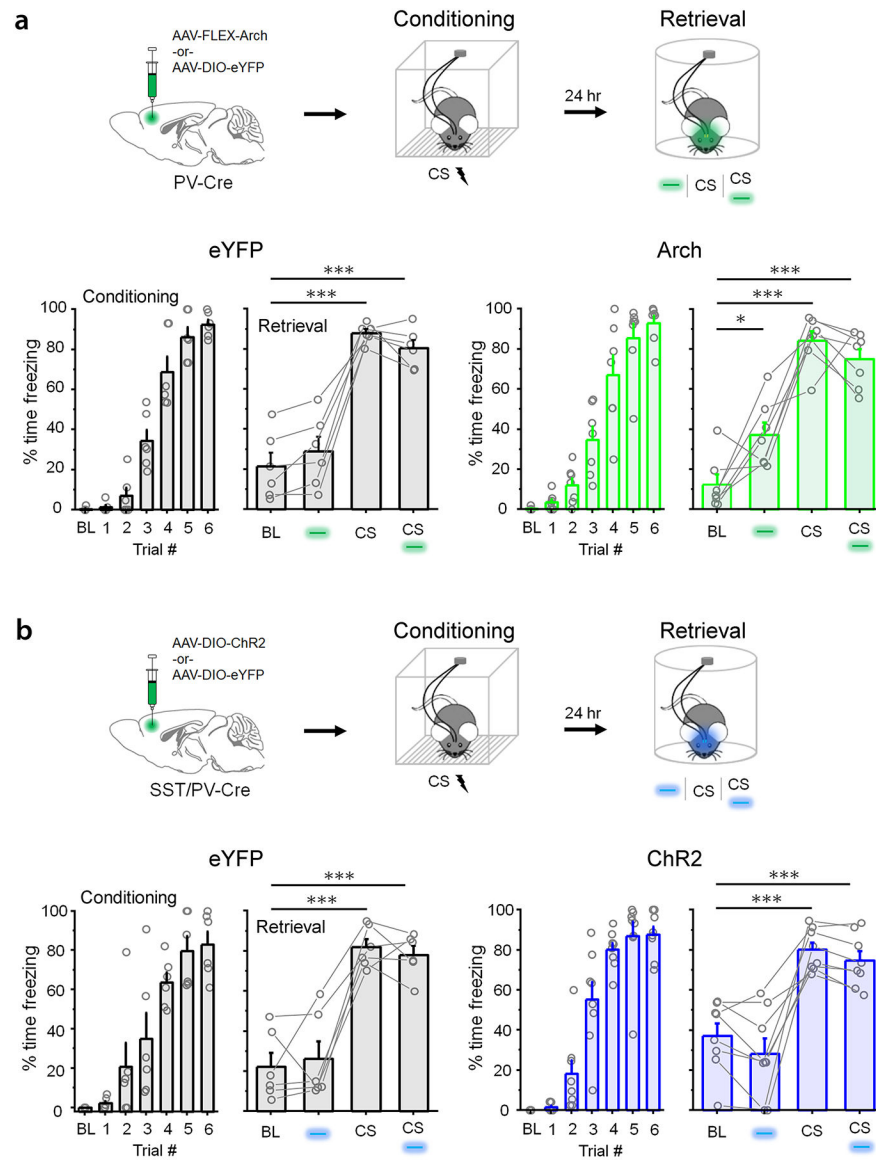


Fig. 6. SST-IN-evoked freezing requires suppression of PV-INs.

a-b. *In vivo* dissection of PV-IN contributions to stimulus-dependent freezing. **a.** Modulation of freezing by CS and light (532 nm, constant, 20 s) at 24 hrs after CS-US pairing in PV-IRES-Cre mice expressing conditional Arch (green) or eYFP control vectors (black). Arch: $F_{3,18} = 33.2$, $p = 1.52 \times 10^{-7}$, 1-way repeated measures ANOVA; $n = 6$ mice. eYFP: $F_{3,15} = 50.1$, $p = 4.74 \times 10^{-8}$, 1-way repeated measures ANOVA; $n = 7$ mice. Experiment was performed in two independent cohorts and mice were pooled together for analysis. **b.** Modulation of freezing by CS and light (473 nm, 10 ms pulses, 20 Hz) at 24 hrs after fear conditioning in SST-IRES-Cre/ PV-IRES-Cre double transgenic mice expressing conditional ChR2 (blue) or eYFP control vectors (black). ChR2: $F_{3,21} = 23.3$, $p = 6.94 \times 10^{-7}$, 1-way repeated measures ANOVA; $n = 8$ mice. eYFP: $F_{3,15} = 41.3$, $p = 1.74 \times 10^{-7}$, 1-way repeated measures ANOVA; $n = 6$ mice. Experiment was performed in two independent

cohorts and mice were pooled together for analysis. * $p < 0.05$, *** $p < 0.001$ by Tukey's post-hoc test. Bar graphs depict mean \pm SE.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

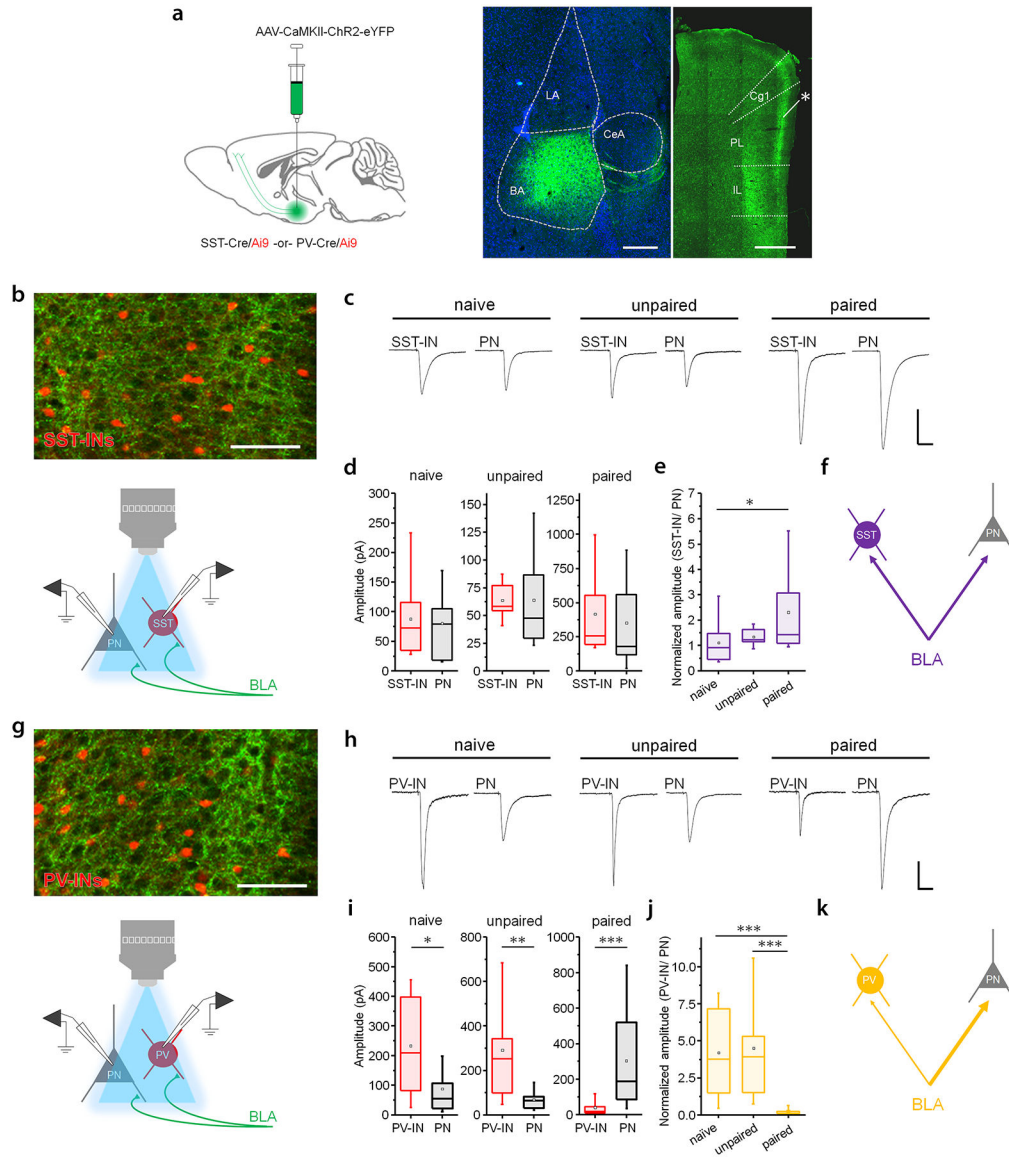


Fig. 7. Relative strength of BLA transmission onto prelimbic INs and PNs is modulated by learning.

a, To examine prelimbic (PL) neuronal responses to BLA afferent stimulation, a CaMKII promoter-dependent ChR2 vector was infused into the BLA (left) of SST- or PV-IRES-Cre/Ai9 mice, leading to projection-specific expression of ChR2-eYFP in PL L2/3 (right). LA = lateral amygdala, BA = basal amygdala, CeA = central amygdala, Cg1 = cingulate area 1, IL = infralimbic. * PL ChR2 accumulation. Scale = 250 (left) and 500 (right) μm. At 24 hours after CS-US pairing, recordings were obtained from Tomato⁺ INs as well as surrounding PNs during BLA afferent photoexcitation (460 nm, 1 ms pulse, 0.1 Hz). **b**, Axonal ChR2 accumulation in layer 2/3 of SST-IRES-Cre/Ai9 mice and recording configuration. **c**, Example EPSC traces. Scale = 100 pA × 50 ms. **d**, Amplitude of BLA-evoked SST-IN and PN EPSCs in slices from naïve mice (8 slices from 4 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 134$, $p = 0.51$, two-sided Mann-Whitney U test; SST-IN, $n = 18$ cells; PN, $n =$

13 cells. Effect of cell type (unpaired): $U = 106$, $p = 0.49$, two-sided Mann-Whitney U test; SST-IN, $n = 13$ cells, PN, $n = 14$ cells. Effect of cell type (paired): $U = 107$, $p = 0.37$, two-sided Mann-Whitney U test; PV-IN, $n = 16$ cells; PN, $n = 11$ cells. **e**, Amplitude of EPSCs in SST-INs normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 7.2$ (2), $p = 0.027$, Kruskal-Wallis ANOVA. **f**, Relative strength of BLA transmission in conditioned mice. **g**, Axonal ChR2 accumulation in layer 2/3 of PV-IRES-Cre/ Ai9 mice. **h**, Example EPSC traces. Scale = $100 \text{ pA} \times 50 \text{ ms}$. **i**, Amplitude of BLA-evoked PV-IN and PN EPSCs in slices from naïve mice (6 slices from 3 mice), as well as those that received unpaired (6 slices from 3 mice) or paired training (8 slices from 4 mice). Effect of cell type (naïve): $U = 118$, $p = 0.025$, two-sided Mann-Whitney U test; PV-IN, $n = 14$ cells; PN, $n = 11$ cells. Effect of cell type (unpaired): $U = 85$, $p = 0.0066$, two-sided Mann-Whitney U test; PV-IN, $n = 10$ cells, PN, $n = 10$ cells. Effect of cell type (paired): $U = 19$, $p = 7.56 \times 10^{-4}$, two-sided Mann-Whitney U test; PV-IN, $n = 13$ cells; PN, $n = 12$ cells. **j**, Amplitude of EPSCs in PV-INs normalized to median values from PNs in the same slices. Effect of training: $\chi^2 = 21.6$ (2), $p = 2.04 \times 10^{-5}$, Kruskal-Wallis ANOVA. **k**, Relative strength of SST-IN transmission in conditioned mice. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by two-sided Mann-Whitney U test (**i**) or Dunn's post-hoc test (**e**, **j**). Box plots depict median (center line), mean (black box), quartiles, and 10–90% range (whiskers).

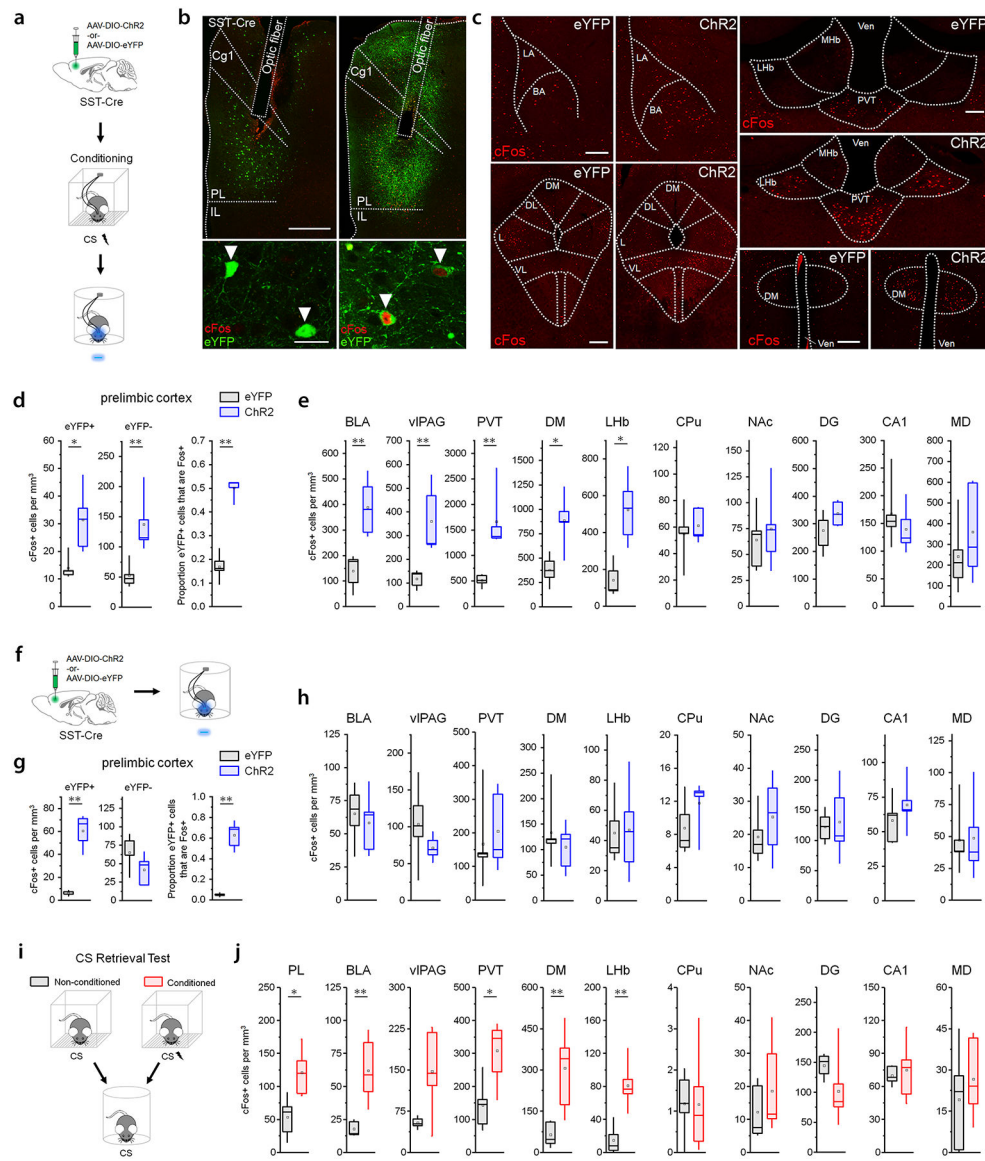


Fig. 8. SST-IN activation recruits a specific brain network in conditioned mice.

a, Preparation of Chr2 ($n = 5$) and eYFP mice ($n = 5$) used for cFos activity mapping after SST-IN photoexcitation. All subjects underwent CS-US pairing 24 hrs prior to stimulation. **b**, Induction of cFos by photoexcitation (6 trials, 473 nm, 10 ms pulses, 20 Hz) in prelimbic (PL) cortex. Insets (bottom) depict eYFP/ cFos double-labeled SST-INs. Arrowheads denote SST-INs. Scale = 500, 50 μm . **c**, Elevated cFos expression in select brain regions after optic stimulation of conditioned mice. LA = lateral, BA = basal, amygdala; DM = dorsomedial hypothalamus; DL = dorsolateral, L = lateral, VL = ventrolateral, periaqueductal gray; MHb = medial habenula; LHb = lateral habenula; PVT = paraventricular thalamus. Scale = 200 μm . **d**, Comparison of cFos⁺ cell counts in PL of stimulated ChR2-eYFP (blue; $n = 5$ mice) versus eYFP-only (black, $n = 5$ mice) mice. In both groups, eYFP⁺ cells represent transfected SST-INs, while eYFP⁻ cells represent non-expressing neighbors. Group differences established by two-sided Mann-Whitney U test and controlled for false discovery

(type 1 error) by the method of Benjamini and Hochberg. eYFP+ cells: $U = 1$, $p = 0.016$. eYFP- cells: $U = 0$, $p = 0.0079$. Proportion eYFP+ cells that are cFos+: $U = 0$, $p = 0.0079$. **e**, Comparison of cFos+ cell counts for all brain regions ($n = 5$ mice per eYFP or Chr2 group, 2 slices per brain region) using the same statistical procedures (two-sided Mann-Whitney U test) as in (c). BLA: $U = 0$, $p = 0.0079$. vIPAG: $U = 0$, $p = 0.0079$. PVT: $U = 0$, $p = 0.0079$. DM: $U = 1$, $p = 0.0022$. LHb: $U = 0$, $p = 0.0012$. Caudate putamen (CPu): $U = 14$, $p = 0.84$. Nucleus accumbens (NAc): $U = 10$, $p = 0.69$. Dentate gyrus of the dorsal hippocampus (DG): $U = 7$, $p = 0.29$. Cornu ammonis area 1 of the ventral hippocampus (CA1): $U = 16$, $p = 0.55$. Mediodorsal thalamus (MD): $U = 8$, $p = 0.42$. Experiment was conducted in two independent cohorts and results were pooled together. cFos counts for each mouse are reported as the average of two representative brain slices per mouse. Tissue quantified in b-e is from the same mice. **f**, Preparation of Chr2 ($n = 5$) and eYFP mice ($n = 5$ mice) used for cFos activity mapping after SST-IN photoexcitation in naïve (unconditioned) mice. **g**, Comparison of cFos+ cell counts in PL of stimulated Chr2-eYFP (blue, $n = 5$ mice) versus eYFP-only (black, $n = 5$ mice) naïve mice (two-sided Mann-Whitney U test). eYFP+ cells: $U = 0$, $p = 0.0079$. eYFP- cells: $U = 20$, $p = 0.15$. Proportion eYFP+ cells that are cFos+: $U = 0$, $p = 0.0079$. **h**, Comparison of cFos+ cell counts in select brain regions of stimulated Chr2-eYFP (blue, $n = 5$ mice) and eYFP-only (black, $n = 5$ mice) naïve mice, two-sided Mann-Whitney U test. BLA: $U = 14$, $p = 0.84$. vIPAG: $U = 19$, $p = 0.22$. PVT: $U = 11$, $p = 0.84$. DM: $U = 13$, $p = 1$. LHb: $U = 13$, $p = 1$. CPu: $U = 7$, $p = 0.31$. NAc: $U = 9$, $p = 0.55$. DG: $U = 12$, $p = 1$. CA1: $U = 6$, $p = 0.22$. MD: $U = 14$, $p = 0.84$. Experiment was conducted in 2 independent cohorts and results were pooled together. cFos counts for each mouse are reported as the average of two representative brain slices per mouse. Tissue quantified in g-h is from the same mice. **i**, Preparation of SST-IRES-Cre mice used for cFos activity mapping after CS exposure (4 trials, 2 KHz, 80 dB, 20 s) in mice that were conditioned 24 hours prior ($n = 6$) or in those that did not receive conditioning ($n = 5$). **j**, Comparison of cFos+ cell counts in select brain regions of conditioned ($n = 6$ mice) versus nonconditioned ($n = 5$ mice) mice, as conducted for photostimulated mice (**e** and **h**). PL: $U = 2$, $p = 0.0017$. BLA: $U = 0$, $p = 0.0043$. vIPAG: $U = 5$, $p = 0.082$. PVT: $U = 3$, $p = 0.030$. DM: $U = 0$, $p = 0.0043$. LHb: $U = 0$, $p = 0.0043$. CPu: $U = 17$, $p = 0.79$. NAc: $U = 9$, $p = 0.33$. DG: $U = 25$, $p = 0.082$. CA1: $U = 12$, $p = 0.66$. MD: $U = 12$, $p = 0.66$. Experiment was conducted in 2 independent cohorts and results were pooled together. cFos counts for each mouse are reported as the average of two representative brain slices per mouse. Tissue quantified in i-j is from the same mice. * $p < 0.05$, ** $p < 0.01$ by two-sided Mann-Whitney U test. Box plots depict median (center line), mean (black box), quartiles, and 10–90% range (whiskers).