HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2023 September 21.

Published in final edited form as:

Neuron. 2022 September 21; 110(18): 3036–3052.e5. doi:10.1016/j.neuron.2022.07.004.

Control of fear by discrete prefrontal GABAergic populations encoding valence-specific information

Kirstie A. Cummings[2,*], **Sabina Bayshtok**[1], **Tri N. Dong**[1,4], **Paul J. Kenny**[1,3], **Roger L. Clem**[1,*,†]

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

[2]Department of Neurobiology, University of Alabama at Birmingham School of Medicine, Birmingham, AL

[3]Drug Discovery Institute, Icahn School of Medicine at Mount Sinai, New York, NY

[4]Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

Summary

Neurons activated by learning have been ascribed the unique potential to encode memory, but the functional contribution of discrete cell types remains poorly understood. In particular, it is unclear whether learning engages specific GABAergic interneurons and, if so, whether they differ functionally from interneurons recruited by other experiences. Here we show that fear conditioning activates a heterogeneous neuronal population in the medial prefrontal cortex (mPFC) that is largely comprised of somatostatin-expressing interneurons (SST-INs). Using intersectional genetic approaches, we demonstrate that fear learning-activated SST-INs exhibit distinct circuit properties and are selectively reactivated to mediate cue-evoked memory expression. In contrast, an orthogonal population of SST-INs activated by morphine experience exerts opposing control over fear and supports reward-like motivational effects. These results outline an important role for discrete subsets of GABAergic cells in emotional learning and point to an unappreciated capacity for functional specialization among SST-INs.

Graphical Abstract

[*]Correspondence should be addressed to: roger.clem@mssm.edu, kac3@uab.edu. [†]Lead contact: roger.clem@mssm.edu.

Author contributions

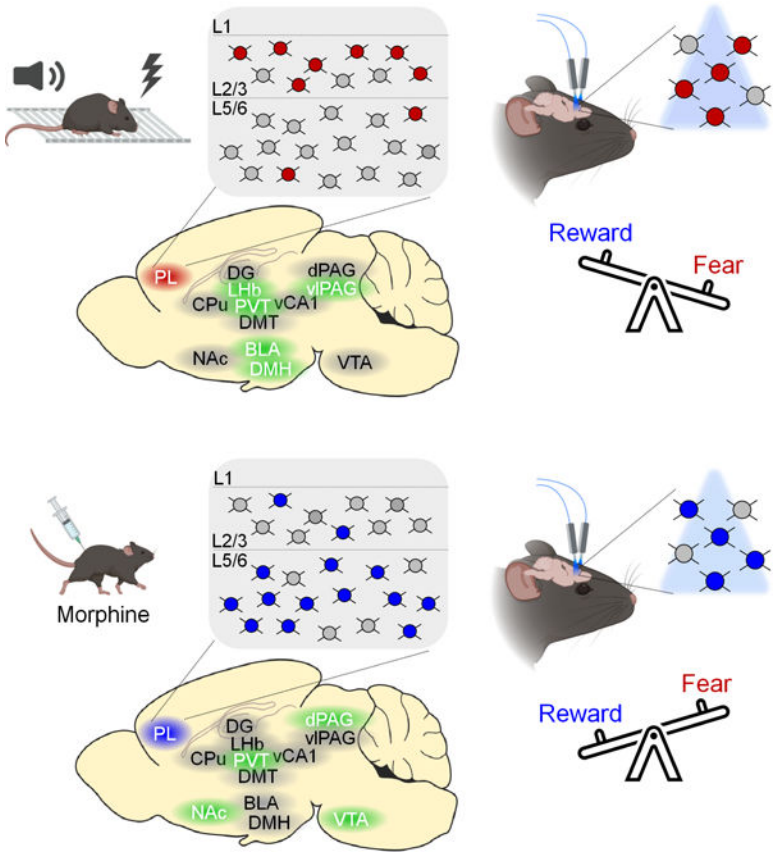
K.A.C. and R.L.C. initiated the project; R.L.C. and P.J.K. supervised research; K.A.C. and R.L.C. designed experiments; K.A.C., S.B. and K.N.D. performed the research and data analysis; R.L.C. and K.A.C. wrote the manuscript.

Declaration of interests

P.J.K. is co-founder and shareholder in Eolas Therapeutics Inc., which has a licensing agreement with AstraZeneca to develop a novel therapeutic for the treatment of substance use disorders.

List of supplementary materials

Supplemental Figures 1–8 and legends



eTOC blurb

Cummings et al. demonstrate that discrete populations of prefrontal somatostatin interneurons activated by positive and negative experience exert opposing control over fear memory. This outlines an important role for functional specialization of GABAergic interneurons in emotional learning.

Keywords

Somatostatin; interneurons; medial prefrontal cortex; Pavlovian; morphine; memory; engram

Introduction

Stimuli encountered during learning preferentially activate specific neurons. Studies that monitor and manipulate these populations suggest that they play an essential role in memory encoding and may retain a cellular trace of learning that, when reactivated, mediates memory expression (Bocchio et al., 2017; Clem and Schiller, 2016; Josselyn and Tonegawa, 2020). Despite the theoretical importance of these stimulus representations, the cellular composition of memory-related populations is largely unexplored, and therefore the functional contribution of specific cell types, defined by gene expression, morphology, or input-output connectivity, remains poorly understood.

The formation of associative fear memory is primarily attributed to glutamatergic projection neurons (PNs), which express a variety of cellular mechanisms for experience-dependent plasticity (Herry and Johansen, 2014; Malenka and Bear, 2004; Ressler and Maren, 2019). However, GABAergic INs also respond to memory-related cues and extensively modulate the function of PNs. In particular, INs containing parvalbumin- (PV-INs), vasoactive intestinal peptide (VIP-INs) and somatostatin (SST-INs) play key roles in orchestrating circuit dynamics underlying memory acquisition and expression (Artinian and Lacaille, 2018; Courtin et al., 2014; Cummings and Clem, 2020; Krabbe et al., 2019; Lucas and Clem, 2017; Wolff et al., 2014; Xu et al., 2019). Synaptic inhibition is thought to facilitate these processes in part through rhythmic entrainment of PN firing (Headley and Pare, 2017). However, evidence suggests that both learning and recall also rely on synaptic interactions between INs that promote PN firing through disinhibition, for example when the recruitment of one IN subtype leads to a corresponding suppression of another (Artinian and Lacaille, 2018; Letzkus et al., 2015; Lucas and Clem, 2018). Because they modulate large groups of PNs, such complex outcomes of GABAergic transmission could endow discrete subsets of INs with unique influence over memory networks. Consequently, an important unanswered question is whether learning recruits specific INs that participate selectively in memory encoding.

Recently, we demonstrated that SST-INs in the prelimbic cortex exhibit increased synaptic efficacy as well as cue-evoked activity after auditory fear conditioning, suggesting their involvement in memory formation (Cummings and Clem, 2020). We further showed that SST-INs control defensive freezing through activation of a distributed brain network, an effect likely orchestrated through the relief of PV-IN-mediated inhibition. Here, using genetic tagging approaches, we investigate whether SST-INs function in memory encoding as part of a discrete cell population. Consistent with a stable cellular representation, prelimbic neurons activated by fear conditioning are essential for subsequent memory expression. Contributing prominently to this heterogeneous population are a distinct subset of SST-INs that are preferentially reactivated at memory retrieval and underlie defensive freezing. Conversely, cellular tagging during morphine treatment, a primarily rewarding experience (Jiang et al., 2021; Le Merrer et al., 2009), delineates an orthogonal population of SST-INs that is not only distinct from fear-related SST-INs, but exerts opposing control over fear expression.

Results

Genetic capture of memory-related prefrontal neurons.

Previous work has identified the dorsomedial mPFC, comprised largely of the prelimbic cortex, as an important site for conditioned stimulus (CS) processing in auditory fear conditioning and a critical substrate for fear expression (Bagur et al., 2021; Burgos-Robles et al., 2009; Corcoran and Quirk, 2007; Courtin *et al.*, 2014; Cummings and Clem, 2020; Dejean et al., 2016; DeNardo et al., 2019; Herry and Johansen, 2014; Karalis et al., 2016; Sotres-Bayon and Quirk, 2010; Xu *et al.*, 2019). To identify neurons selectively recruited by learning in the prelimbic mPFC, we employed a viral genetic approach to permanently tag these cells with a fluorescent reporter. This strategy relied on expression of an estrogen

receptor-dependent Cre-recombinase (ERCreER) under the control of the enhanced synaptic activity responsive element (E-SARE), which contains regulatory sequences derived from the *arc* immediate-early gene (Kawashima et al., 2013). Importantly, recombination of target alleles following induction of ERCreER depends on the presence of the estrogen receptor ligand 4-hydroxytamoxifen (4-OHT), which together with rapid degradation of ERCreER restricts neuronal tagging to a period of several hours following activity.

A viral vector containing the E-SARE-ERCreER construct was co-injected into the mPFC together with vectors encoding a Cre-dependent enhanced yellow fluorescent protein (DIO-eYFP) and synapsin-driven mCherry (hSyn-mCherry), the latter of which served purely as a marker of viral infusion (Figure 1A). Following viral incubation, mice were subjected to auditory fear conditioning, which consisted of paired presentations of an auditory tone (2 KHz, 80 dB, 20s) that co-terminated with an aversive foot shock unconditioned stimulus. As a behavioral control condition, a subset of animals underwent the same procedure, except that aversive foot shocks were omitted (tones only). Following conditioning or tone presentations, subjects received intraperitoneal injections of 4-OHT or vehicle (Veh) solution and were returned to their home cages for 2 weeks to allow for the expression of recombinant alleles. At this point they were subjected to a test of memory retrieval in a context distinct from the training arena (Figure S1), followed by cFos immunohistochemistry, to examine the degree to which neurons that were activated during CS-US pairing were specifically reactivated by the CS.

Mice injected with 4-OHT after fear conditioning, but not after tones only exposure, exhibited substantially more eYFP+ neurons than vehicle-treated mice, which contained only sparse background labeling (Figure 1B, C). Furthermore, while no overall differences in cFos levels were observed following memory retrieval (Figure 1D), mice that received 4-OHT following fear conditioning exhibited a higher density of double-labeled neurons (Figure 1E). More importantly, a higher proportion of eYFP+ neurons in these animals were immunoreactive for cFos (Figure 1F), indicating that a greater incidence of double labeling was not attributable to higher levels of eYFP expression. This indicates that neurons activated during fear conditioning are preferentially reactivated during memory retrieval, consistent with their selective involvement in long-term memory.

To test whether learning-activated neurons in the prelimbic cortex mediate memory expression, we next examined whether optogenetic manipulation of these cells influences CS-evoked freezing. To enable optogenetic control of these populations, mice received bilateral infusions of a viral cocktail including E-SARE-ERCreER and hSyn-mCherry vectors, along with a vector encoding either a Cre-dependent Archaerhodopsin (FLEX-Arch3.0-GFP; Arch) or channelrhodopsin construct (DIO-ChR2-eYFP; ChR2) and were implanted with optic fibers directed at prelimbic cortex (Figure S2). After CS-US pairing, subjects received intraperitoneal injections of either 4-OHT or vehicle and were submitted to a memory retrieval test in which we examined the independent and combined effects of light and CS trials on freezing behavior. Following expression of Arch within learning-activated neurons, freezing during CS trials combined with photoinhibition (532 nm, 20 s epoch, constant) was indistinguishable from baseline (Figure S2). Conversely, after ChR2 expression, photoexcitation (473 nm, 20 s epoch, 20 Hz, 5 ms duration)

increased freezing even in the absence of the CS. No further increase in freezing was observed when photoexcitation was combined with CS presentation, which could potentially reflect a ceiling effect. Examination of difference scores revealed bidirectional effects of optogenetic manipulations during the baseline period, suggesting that a component of the tagged population might be involved in context fear. The same analysis performed for CS trials indicated a decrease in freezing in 4-OHT-injected Arch-expressing animals compared to all other groups, supporting a role for the tagged population in cued freezing. Importantly, photostimulation did not affect general locomotion in the open field (Figure S2). Furthermore, no light-dependent effects were observed in the vehicle-injected groups, or in animals that received 4-OHT after tones only exposure (Figure S2, S3), indicating that modulation of freezing requires prior conditioning in conjunction with ERCreER activity.

Collectively, PV-INs, VIP-INs and SST-INs comprise >80% of GABAergic neurons in the cortex (Rudy et al., 2011). To determine whether these cell types participate as part of the learning-related population, we performed immunohistochemical staining against markers of these IN subtypes and quantified overlap with an activity-dependent fluorescent tag (Figure 1G-I). In animals that received fear conditioning in combination with 4-OHT, we observed a far greater density of eYFP+ SST-INs compared to all other conditions (Figure 1G). In addition, a higher proportion of tagged neurons in this group expressed SST, indicating that an increase in labeled SST-INs after conditioning cannot be attributed solely to higher eYFP levels. Importantly, these effects were not observed after tones only exposure, further indicating that SST-IN labeling is specific to CS-US pairing and not driven by incidental activity. On average, ~20% of the tagged population was comprised of SST-INs, which in turn represented ~30% of the total SST-IN population (Figure 1G).

In contrast to the above results, learning-activated neurons exhibited negligible immunoreactivity for PV or VIP, and the incidence of recombination among PV- and VIP-INs was not modulated by conditioning or 4-OHT (Figure 1H-I). This could potentially indicate that fear conditioning does not strongly activate these cell types or, alternatively, they may exhibit molecular differences that preclude activation and binding of transcription factors CREB, MEF2 and SRF to E-SARE in response to behavioral experience (Kawashima et al., 2014). While both SST- and VIP-INs respond to aversive foot shocks (Cummings and Clem, 2020; Krabbe *et al.*, 2019; Pi et al., 2013), genetic tagging would require coupling between this type of activity and the above transcriptional responses, which have been linked to memory encoding (Alberini, 2009). Therefore, E-SARE induction potentially demarcates a subset of SST-INs specifically involved in trace formation.

SST-INs activated during fear learning participate in memory encoding.

The above results raise the possibility that SST-INs are among the prefrontal neurons activated by learning and subsequently reactivated to mediate memory retrieval (Figure 1), which are important criteria for engram-bearing cells (Tonegawa et al., 2015). Alternatively, activated SST-INs may be relevant only to initial learning, with recall being mediated by other cell types (e.g. glutamatergic PNs). To discriminate between these possibilities, we made use of intersectional activity-dependent tagging to examine the specific role of learning-activated SST-INs. This strategy resembled the one used for non-selective tagging

except that to be activated by ERCreER, target vectors were also required to undergo flippase (Flp) recombination, which was restricted to SST-INs by the use of SST-FlpO transgenic mice (Figure 2A-B). Prior to training, these animals received bilateral mPFC injections of a viral cocktail including E-SARE-ERCreER and hSyn-mCherry vectors, along with a Cre- and Flp-dependent eYFP construct (Cre_{on}/Flp_{on}-eYFP) (Figure 2B). Subjects then underwent auditory fear conditioning, or tones only experience (Figure S4), and immediately afterwards received intraperitoneal injections of 4-OHT or vehicle. Following a test of memory retrieval in a context distinct from the training arena, we utilized cFos immunohistochemistry to examine the degree to which SST-INs that were activated during CS-US pairing were reactivated during CS exposure.

Mice that received 4-OHT after fear conditioning, but not tones only exposure, exhibited a higher density of eYFP+ cells than vehicle controls (Figure 2C,D). Importantly, we found that a vast majority (>91%) of neurons labeled in this manner are immunoreactive for SST, indicating that our approach captures a relatively pure population of activated SST-INs (Figure S4). Following the test of memory retrieval, comparable levels of cFos expression were exhibited across conditions (Figure 2E). However, conditioning combined with 4-OHT was associated with a higher density of cFos+ eYFP cells following retrieval (Figure 2F) and, importantly, a higher proportion of eYFP+ cells in these animals exhibited cFos immunoreactivity (Figure 2G). This pattern of results suggests that cellular activity during memory retrieval is extraordinarily selective for SST-INs that were active during learning. Indeed, nearly 30% of the tagged SST-IN population was reactivated upon retrieval.

Given these results, we next sought to determine whether SST-INs activated during fear conditioning exert control over memory expression. Prior to training, we injected into the mPFC of SST-FlpO transgenic mice a viral cocktail including E-SARE-ERCreER and hSyn-mCherry vectors, along with a Cre- and Flp-dependent (Cre_{on}/Flp_{on}) Arch or ChR2 construct (Figure 3A-B). Following auditory fear conditioning, or tones only experience, mice received 4-OHT or vehicle solution and were submitted to a test of memory retrieval in which we examined the independent and combined effects of light and CS trials. Strikingly, selective photoinhibition (532 nm, 20 s epoch, constant) of a relatively sparse population of tagged SST-INs abolished CS-evoked freezing (Figure 3C). Conversely, photoexcitation (473 nm, 20 s epochs, 20 Hz, 5 ms pulses) was on its own sufficient to elicit freezing (Figure 3D), an effect similar to that observed during stimulation of a heterogeneous population (Figure S2). Examination of difference scores confirmed a selective effect of photoexcitation in 4-OHT-injected ChR2 animals during the baseline period (Figure 3E), and conversely a selective reduction in freezing in 4-OHT-injected Arch-expressing animals during the CS compared to all other groups. These results were not attributable to non-specific motor effects because photostimulation did not alter locomotor parameters in the open field test (Figure S5). In contrast to these results, mice that received tones only experience or unpaired conditioning prior to SST-IN tagging exhibited no change in freezing upon photoexcitation (Figure S5). These results suggest that fear conditioning activates a subset (~30%) of SST-INs that contribute to encoding of cue associations.

Learning-activated SST-INs exhibit unique input and output synaptic transmission.

An influential hypothesis is that neurons engaged by learning become functionally differentiated from non-activated cells through plasticity. Thus, an important question is whether fear learning-activated SST-INs exhibit differences in input or output synaptic transmission that might explain their preferential reactivation during memory retrieval or their downstream effects within the network. To examine this possibility, we first employed an activity-dependent fluorescent tag in conjunction with constitutive labeling of SST-INs, allowing us to compare the synaptic properties of tagged and non-tagged SST-INs. First, SST-FlpO x Ai65 transgenic mice received bilateral mPFC infusions of E-SARE-ERCreER and Cre_{on}/Flp_{on}-eYFP vectors and underwent fear conditioning followed by 4-OHT injection (Figure 4A). Electrophysiological recordings were then obtained from learning-activated SST-INs, which were labeled with both eYFP and tdTomato, as well as non-tagged SST-INs, which expressed only tdTomato (Figure S9). SST-INs activated during learning exhibited a higher frequency of spontaneous excitatory postsynaptic currents (EPSCs) compared to non-tagged SST-INs (Figure 4B). However, there were no differences in either frequency or amplitude of inhibitory postsynaptic currents (IPSCs) (Figure 4C). Electrically-evoked EPSCs from tagged SST-INs also exhibited a lower paired-pulse ratio (PPR) compared to those from non-tagged SST-INs, indicating that elevated excitatory transmission onto learning-activated SST-INs is attributable to higher glutamate release probability (Figure 4D). These differences were not observed when comparing tagged and non-tagged SST-INs from animals that received unpaired training (Figure S6), suggesting that they are a specific property of cells engaged in cued associative learning, and given our prior findings, they are likely acquired as part of the memory encoding process (Cummings and Clem, 2020). Interestingly, when we examined action potential firing in response to current injection, we also observed higher excitability of tagged versus non-tagged SST-INs after CS-US pairing specifically in male mice (Figure S6).

While increased excitatory input may play a causal role in the reactivation of learning-activated SST-INs, the ability of SST-INs to modulate freezing depends on interaction with other cell types within the mPFC. We previously observed that CS-US pairing, but not unpaired training, alters the balance of transmission from prelimbic SST-INs onto PV-INs versus PNs in a manner that favors PN disinhibition (Cummings and Clem, 2020). We therefore sought to determine whether the microcircuit properties of SST-INs specifically engaged by CS-US pairing differ from those of SST-INs activated under non-associative, or unpaired, conditions. Specifically, we examined the relative level of GABAergic inhibition that these populations provide onto neighboring PV-INs as well as PNs that were either tagged or non-tagged as a result of training, in order to reveal how SST-INs interact selectively with learning-activated PNs. To accomplish this, we first generated SST-FlpO x PV-Cre x Ai9 transgenic mice, and then bilaterally infused a viral cocktail containing E-SARE-ERCreER and Cre_{on}/Flp_{on}-ChR2-eYFP vectors into the mPFC of these animals (Figure 4E). Following fear conditioning and 4-OHT injection, this permitted expression of ChR2-eYFP and tdTomato expression in learning-activated SST-INs and PNs, respectively. In addition, constitutive expression of tdTomato was also driven via the PV-Cre allele within PV-INs, which can be readily distinguished from PNs by their electrophysiological properties.

In acute brain slices, we recorded monosynaptic IPSCs in the above cell populations in response to photoexcitation of learning-activated SST-INs (Figure 4F). In unpaired mice, responses of tagged and non-tagged PNs were similar in amplitude, and both cell types exhibited larger responses than PV-INs (Figure 4F-G). In contrast, in paired mice, response amplitudes in tagged PNs were smaller than in non-tagged PNs, while those in tagged PNs and PV-INs were equivalent. To control for potential differences in viral transduction between paired and unpaired animals, IPSCs from PV-INs and tagged PNs were normalized to responses from non-tagged PNs within the same animal (Figure 4H). This revealed that compared to unpaired training, SST-INs activated during CS-US pairing provide proportionately stronger input onto PV-INs and generate weaker responses within tagged PNs. These differences in output, which potentially result from learning-induced plasticity, may facilitate selective disinhibition of memory-related PNs through concerted mono- and disynaptic control (Figure 4I).

Morphine and fear conditioning activate functionally discrete SST-IN populations.

While the above results suggest that memory-encoding SST-INs have distinct circuit properties, it remains possible that freezing elicited by optogenetic stimulation is a non-specific effect of SST-IN transmission, rather than dependent on reactivation of specific cells. In this case, similar effects would be obtained by stimulating other subsets of SST-INs as long as they exceed some threshold population size. Previous work has shown that prelimbic SST-INs exhibit plasticity of membrane and synaptic properties following morphine treatment, an experience highly distinct from fear conditioning (Jiang *et al.*, 2021). We therefore used activity-dependent genetic tagging to investigate whether SST-INs activated by morphine exhibit overlap with those involved in fear conditioning, and whether reactivation of these cells can drive fear expression after CS-US pairing (Figure 5). SST-FlpO transgenic mice received viral infusions identical to those employed for fear-related cellular tagging (Figure 4) and were then injected with morphine (10 mg/kg) or saline solution followed 10 hours later by 4-OHT or vehicle (Figure 5A). Morphine treated mice exhibited stereotypical hyperlocomotion compared to vehicle injected controls (Figure S7) (Lee et al., 1977). Three weeks later, all mice underwent CS-US pairing followed by CS presentation and immunohistochemical staining for cFos to determine the degree to which morphine-activated SST-INs were reactivated by fear memory retrieval (Figure S7). As expected, a higher density of eYFP-positive SST-INs was observed following morphine combined with 4-OHT, relative to control conditions (Figure 5B-C). Among the groups, however, a comparable proportion of eYFP-positive SST-INs were immunoreactive for cFos, indicating chance levels of overlap between morphine-related tagging and retrieval-induced activity (Figure 5D-F). This indicates that in contrast to SST-INs activated during fear conditioning (Figure 4), morphine-activated SST-INs were not preferentially reactivated by memory retrieval.

To examine the impact of morphine-activated SST-INs on fear expression, cellular tagging was performed as above, except that E-SARE-ERCreER was used to drive Arch- or ChR2-eYFP expression to enable subsequent optogenetic manipulations (Figure 5G). Three weeks later, neither photoinhibition (532 nm, 20 s epochs, constant) nor photophotoexcitation of morphine tagged SST-INs (473 nm, 20 s epochs, 20 Hz, 5 ms pulses) affected freezing

levels in shock-naïve animals (Figure S7). On the next day, all subjects underwent CS-US pairing followed by a test of fear memory retrieval in which we examined the independent and combined effects of light and CS trials. In contrast to stimulation of SST-INs that were activated during fear conditioning (Figure 3), photoinhibition of morphine-activated SST-INs had no effect on CS-evoked freezing (Figure 5H). Likewise, photoexcitation of morphine-activated SST-INs had no effect on baseline freezing levels (Figure 5I). Interestingly, however, when photoexcitation was combined with CS presentation, freezing levels were reduced compared to CS-only trials (Figure 5I). Importantly, photoexcitation of morphine-activated SST-INs had no effect on open field behavior, implying that reduction of freezing during memory retrieval was not attributable to non-specific locomotor effects (Figure S7). These results indicate that although both morphine and fear conditioning recruit large populations of SST-INs, these neurons have distinct functional properties. Not only is memory retrieval uniquely dependent on SST-INs that were activated by fear conditioning, but morphine-related SST-INs exert opposing control over fear expression.

Morphine- and fear-related SST-INs signal opposite valence and modulate distinct brain networks.

Behavioral effects of activating morphine- and fear-related SST-INs suggest that these populations may support discrete forms of valence signaling. Consistent with this possibility, mPFC plays an important role in opiate reward (Hyman et al., 2006; Rosen et al., 2015). To examine whether SST-INs activated by morphine support reward-like signaling, we tested the motivational effects of optogenetic stimulation in an assay of conditioned place preference (CPP). After establishing an initial baseline preference, mice received three consecutive days of CPP training in which photoexcitation of morphine tagged SST-INs was paired with the compartment opposite the one for which an initial preference was expressed (Figure 6A). Following CPP training, mice that received both morphine and 4-OHT during SST-IN tagging displayed a shift in preference toward the compartment where stimulation was applied, relative to control animals that received vehicle instead of 4-OHT (Figure 6B-F). Because this preference was expressed in the absence of photostimulation, it indicates that activation of morphine-related SST-INs is sufficient to induce the formation of a positive contextual association.

Reactivation of fear-related SST-INs elicits defensive freezing (Figure 3), which on its own indicates that stimulation of these cells is aversive. Nevertheless, to examine the valence properties of fear-related SST-INs under conditions comparable to those in which morphine-related SST-INs elicit reward-like effects, we conducted an assay of conditioned place aversion (CPA). In this case, photoexcitation was paired with the compartment for which the animal expressed an initial preference (Figure 6G). Following CPA training, mice that received photoexcitation of fear-related SST-INs exhibited lower preference for the compartment where stimulation was applied, relative to vehicle controls (Figure 6H-L). These data confirm that under similar conditions, activation of morphine- versus fear-related SST-INs elicits opposite motivational effects and are therefore perceived as having opposite valence.

To gain further insight into functional specialization of fear- and morphine-related SST-INs, we examined the distribution of these cells within the prelimbic cortex as well as their modulation of activity in potential downstream regions. Interestingly, while a greater proportion of fear-related SST-INs could be found in layer 2/3, the distribution of morphine-related SST-INs was highest in layer 5/6 (Figure 7A-C). Following E-SARE-dependent tagging with ChR2 and prelimbic fiber implantation, we applied photoexcitation to determine whether discrete SST-IN populations control activity of brain regions implicated in aversion and reward. After photoexcitation of either fear- or morphine-related populations (6 trials, 473 nm, 20 s duration, 5 ms pulse, 20 Hz), the vast majority of ChR2-eYFP-expressing SST-INs exhibited cFos immunoreactivity regardless of whether animals received 4-OHT or vehicle during intersection tagging (Figure 7D-F). Activation of fear-related SST-INs was associated with cFos expression in areas related to defensive threat responding, including the BLA (Herry and Johansen, 2014), vlPAG (Tovote et al., 2016), PVT (Beas et al., 2018; Do-Monte et al., 2015), LHb (Stamatakis and Stuber, 2012) and DMH (Johnson and Shekhar, 2006) (Figure 7G). With the exception of PVT, however, these areas were not differentially activated following stimulation of morphine-related SST-INs (Figure 7H). Instead, stimulation of morphine-related SST-INs resulted in robust activation of NAc and VTA, which form a circuit critical for reward processing (Hyman *et al.*, 2006). In addition, we observed strong activation of the dPAG, comprised of the dorsomedial and dorsolateral subregions. These results suggest that discrete valence properties of fear- and morphine-related populations may be attributable to their unique anatomical distribution as well as their differential control over prelimbic outputs exerting direct or indirect control over remote brain regions.

Given the unique circuit properties of fear-related SST-INs (Figure 4), we examined whether morphine-activated SST-INs likewise exhibit differences in synaptic input compared to non-activated SST-INs or exert differential control over morphine-related PNs. SST-FlpO x Ai65 transgenic mice received bilateral mPFC infusions of E-SARE-ERCreER and Cre_{on}/Flp_{on}-eYFP vectors and were treated with morphine followed 10 hrs later by 4-OHT injection (Figure 8A). Electrophysiological recordings revealed a higher frequency of spontaneous EPSCs as well as lower amplitude of spontaneous IPSCs in morphine-activated SST-INs compared to non-activated SST-INs (Figure 8B-C). Electrically-evoked EPSCs also exhibited a lower PPR in morphine-activated SST-INs compared to those from non-activated cells. In addition to these synaptic differences, morphine-related SST-INs exhibited higher intrinsic excitability (Figure S8). Optogenetic stimulation of morphine-activated SST-INs, labeled in a manner similar to fear-related SST-INs in SST-FlpO x PV-Cre x Ai9 triple transgenic mice, revealed that they elicit much weaker inhibition in PNs that were activated by morphine treatment than in non-tagged PNs (Figure 8E-G). Thus, circuits formed by morphine- and fear-related SST-INs exhibit a similar pattern of electrophysiological differences, but likely differ in their specificity for recruitment versus inhibition of specific output populations. In either case, PNs that participate as part of an activated subpopulation with SST-INs receive less inhibition from these cells, which may facilitate their selective disinhibition upon SST-IN recruitment.

Discussion

In this study, we demonstrate that a heterogeneous population of neurons in the prelimbic mPFC plays a critical role in fear memory encoding. A major subset of cells activated during learning are SST-INs, whose responsiveness to activity-dependent genetic tagging differentiates them from both PV- and VIP-INs. Intersectional genetic capture of learning-activated SST-INs reveals that they are selectively reactivated during memory retrieval and support the expression of defensive freezing, properties that are attributable to differences in transmission within synaptic networks formed by these cells. In particular, memory-related SST-INs receive stronger excitatory input and exhibit GABAergic output that favors disinhibition of PNs recruited by learning. Importantly, fear encoding is a specific property of these cells that is not generalizable to other SST-IN populations. Indeed, SST-INs activated by a distinct positive experience, morphine exposure, exert opposing control over cued fear expression and modulate a brain network implicated in reward.

Our identification of a sparse cell population with predicted attributes of a fear memory engram is comparable to prior studies that employed genetic tagging of learning-activated neurons (Cai et al., 2016; DeNardo *et al.*, 2019; Han et al., 2009; Kim and Cho, 2017; Lacagnina et al., 2019; Liu et al., 2012; Rashid et al., 2016; Reijmers et al., 2007; Ryan et al., 2015; Tayler et al., 2013). While most reports do not explicitly address the cellular composition of engram populations or focus exclusively on the role of glutamatergic projections, however, we demonstrate that memory is represented to a large degree by GABAergic cells in the prelimbic cortex. Indeed, activation of SST-INs recapitulates the effect of stimulating the broader population, implying that they can largely orchestrate network dynamics underlying memory expression and may control the reactivation of glutamatergic PNs. This underscores the importance of establishing how both excitatory and inhibitory neurons contribute to the formation of a cellular memory trace, as well as how these cell types interact during processing of memory-related cues.

A possible key to the function of memory-encoding SST-INs is that they exert differential control over other cell populations involved in learning. Like those activated by unpaired conditioning, SST-INs activated by CS-US pairing give rise to monosynaptic connections onto PV-INs as well as both activated and non-activated PNs (Figure 4). However, the balance of GABAergic transmission onto these populations differed between the training conditions, reflecting either the recruitment of discrete SST-INs or the expression of plasticity in their output synapses. SST-INs engaged by cued fear learning provide proportionately stronger input onto PV-INs, consistent with previous recordings involving non-selective SST-IN stimulation (Cummings and Clem, 2020). In addition, while SST-INs activated by unpaired training inhibit PNs uniformly, those activated by CS-US pairing inhibit activated PNs more weakly than those that were not activated by learning. This unique configuration may facilitate PV-IN-mediated disinhibition while simultaneously sparing fear-related PNs from direct suppression by SST-INs, thereby enabling recruitment of relevant outputs from the prelimbic cortex while preserving or even augmenting the suppression of irrelevant pathways.

In contrast to SST-INs activated by fear conditioning, those labeled in response to morphine treatment did not support CS-evoked freezing but on the contrary opposed its expression (Figure 5). Involvement of these subpopulations in distinct valence systems may in part explain their conflicting behavioral effects, but it remains to be established to what extent this is a pre-existing property of individual SST-INs. This is an important question given the unique laminar distribution of fear- and morphine-related SST-INs and the likelihood that they modulate distinct anatomical networks through PN disinhibition. In addition, however, valence-specific SST-INs may compete for behavioral control through mutually opposing inhibition of these output populations (Garcia-Junco-Clemente et al., 2017). Future work should therefore establish whether reward-related SST-INs interact directly with prelimbic circuits underlying fear expression.

Finally, while important insights can be gained from parsing the contributions of GABAergic subtypes, functional classification of these cells based on molecular markers is inherently complicated. As one of the most diverse of the broadly defined subtypes, SST-INs have been subclassified based on distinct firing phenotypes, morphologies, and co-expression of secondary markers (Yavorska and Wehr, 2016). For example, while Martinotti cells are estimated to comprise at least half of SST-INs, disinhibition of the thalamorecipient layer of somatosensory cortex is mediated by SST-INs with non-Martinotti morphology (Xu et al., 2013). Within the frontal cortex, functionally relevant subpopulations of SST-INs have been discriminated based on expression of receptors for acetylcholine and oxytocin, as well as signaling components like neuropeptide Y and neuronal nitric oxide synthase (Funk et al., 2017; Jackson et al., 2018; Li et al., 2016; Nakajima et al., 2014; Yamamuro et al., 2020). Nevertheless, our appreciation of how these discrete subpopulations encode different aspects of experience is severely limited. Functional mapping of SST-INs, aided by activity-dependent genetic tagging, may help resolve these questions.

In conclusion, conditioned fear relies on an anatomically and functionally discrete population of SST-INs in the prefrontal cortex but is subject to opposing control by an orthogonal network of reward-related SST-INs. Our results provide important insight into the cellular specificity of memory encoding by GABAergic microcircuits and outline synaptic mechanisms by which INs may become functionally specialized during behavioral plasticity.

STAR Methods

Resource availability

Lead contacts—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contacts, Roger Clem (roger.clem@mssm.edu) and Kirstie Cummings (kac3@uab.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—The data that support the findings of this study are available from the corresponding author upon reasonable request. This paper does not report original code.

Experimental model and subject details

All procedures were approved by the Institutional Animal Care and Use Committee at the Icahn School of Medicine at Mount Sinai. Male and female mice aged P42–45 at the time of surgery and P70-P98 at the time of behavioral and electrophysiological experiments were used throughout this study. Mice (C57Bl/6J background) were purchased from Jackson Laboratories and housed in groups of 3–5 in a 12 hour light-dark cycle with *ad libitum* access to food and water. Behavioral and electrophysiological experiments were performed during the animals' light cycle. The following genotypes were used: C57Bl/6J (Stock No. 000664), SST-FlpO (Stock No. 028579), PV-IRES-Cre (Stock No. 017320), Ai9 (Stock No. 007909), and Ai65F (Stock No. 032864). Animals were randomly assigned to experimental groups.

Method Details

Vectors—Vectors used in this study that were purchased from Addgene include AAV1-EF1a-DIO-hChR2(H134R)-eYFP-WPRE (Addgene # 20298), AAV1-CBA-FLEX-Arch-GFP (Addgene # 22222), AAV1-EF1a-DIO-eYFP-WPRE (Addgene # 27056), and AAV8-hSyn-mCherry (Addgene # 114472). Vectors purchased from the University of North Carolina Gene Therapy Center Vector Core include rAAVDJ/nEF-Cre_{on}/Flp_{on}-hChR2(H134R)-eYFP and rAAVDJ/nEF-Cre_{on}/Flp_{on}-eYFP. The plasmid for the ESARE-ERCreER-PEST vector was a gift from Dr. H. Bito (University of Tokyo). We expanded the plasmid in transformation-competent E. coli, followed by purification using a Qiagen MaxiPrep kit and further extraction in a mixture of phenol-chloroform-isoamyl alcohol (25:24:1) saturated with 10 mM Tris (pH 8.0) and 1 mM EDTA. The purified plasmid was packaged in the AAV8 serotype at the Boston Children's Hospital vector core.

Stereotaxic surgical procedures—Surgery was performed as previously described (Cummings and Clem, 2020). Briefly, following induction of anesthesia, mice (P42) were mounted in stereotaxic frames. For activity-dependent tagging in C57Bl/6J mice, vectors were mixed in a ratio of 2:7:1 for ESARE-ERCreER, DIO-eYFP or DIO-ChR2-eYFP or FLEX-Arch-GFP, and hSyn-mCherry, respectively. For intersectional activity-dependent tagging in SST-FlpO mice, vectors were mixed in a ratio of 2:4:0.6 for ESARE-ERCreER, Cre_{on}/Flp_{on}-ChR2-eYFP or Cre_{on}/Flp_{on}-eYFP, and hSyn-mCherry, respectively. Vectors were thoroughly mixed just prior to surgery and bilaterally infused into prelimbic cortex (400 nL; AP +1.9, DV −2.0, ML +/− 0.9 at a 10° angle) at a rate of 100 nl/min using motorized injectors (World Precision Instruments). For *in vivo* optogenetics experiments, optic fibers were fabricated as previously described (Cummings and Clem) and bilaterally implanted directly above prelimbic cortex at AP: +1.9, DV: −1.6, ML: +/−0.9 at a 10° angle. Virus was allowed in incubate for 4 weeks prior to behavioral training experiments.

Fear conditioning and retrieval behavior—All mice were handled for 3 consecutive days for prior to behavioral testing as described previously (Cummings and Clem, 2020). For optogenetic manipulation experiments, mice were also habituated to the patch cords during handling and were tethered to patch cords for all behavioral tests. Auditory fear conditioning was conducted in sound-attenuating chambers (MedAssociates, St. Albans, VT, USA) and began with a 200 s baseline period followed by presentation of 6 pairings of

a neutral auditory tone (CS; 2 kHz, 80 dB, 20 s) with a co-terminating foot shock (US; 0.7 mA, 2 s). CS/US pairings were presented with 80 s interstimulus intervals. For naïve mice, animals were placed in the conditioning arena and exposed to 6 tones (same as CS in conditioning experiments) in the absence of shocks. Unpaired conditioning was performed by placing animals in the conditioning arena and presenting 6 CS, then returned to their home cage for 15 min, after which they were placed back in the conditioning arena and exposed to 6 US.

Two (for nonintersectional tagging in wildtype mice) or three (for intersectional tagging in SST-FlpO mice) weeks after behavioral training (as described below), mice were subjected to a CS-evoked memory retrieval test in a neutral context. For optogenetic experiments, mice received two laser-only trials, followed by 4 CS presentations alternating with and without simultaneous laser stimulation. The order of CS-only and combined CS/laser trials was counterbalanced. Laser-only, CS/laser, and CS-only epochs were each averaged and reported. For trials including optogenetic manipulation of tagged neurons, light was delivered at a final transmitted intensity of 7–9 mW for 473 nm (for ChR2; 20 Hz, 5 ms for 20 s epochs) or 532 nm (for Arch; constant light, 20 s) using a diode-pumped solid state laser (Opto Engine, LLC, Midvale, UT, USA). For mice tethered to an optic fiber, video recordings of fear conditioning and retrieval behavior were scored manually by an experimenter blind to the group allocation of the subject. For untethered mice (i.e. for cFos experiments and electrophysiology), videos were submitted to automated analysis by MedAssociates VideoFreeze software. Mice with mistargeted virus and/or optic fibers were excluded from analysis.

Morphine administration—Morphine sulfate was dissolved in sterile saline at 1 mg/mL. Mice received a single 10 mg/kg intraperitoneal injection of morphine or saline in the home cage. Mice receiving an injection of morphine displayed characteristic phenotypes, including morphine-induced tail erection and hyperlocomotion (Lee *et al.*, 1977), and were left undisturbed for 10 hours prior to receiving injections of vehicle or 4-hydroxytamoxifen (see below).

Activity-dependent neural tagging—Population tagging was performed via a single intraperitoneal (IP) injection of either vehicle or 4-hydroxytamoxifen (4-OHT) immediately following CS-US pairing, tone-only exposure, or foot shock in unpaired mice, or 10 hours following morphine or saline injection. 4-OHT was formulated as previously described (Ye et al., 2016). Briefly, 4-OHT (Sigma H6278) was dissolved in DMSO to a concentration of 40 mg/mL, and further diluted in sterile saline containing 1% Tween-80 to a concentration of 1 mg/mL. Mice were injected at a dose of 10 mg/kg. Vehicle consisted of the same components as the 4-OHT solution, including 2.5% DMSO and 1% Tween-80 in sterile saline, but lacked 4-OHT. Vehicle was injected at the same volume as the 4-OHT mixture (0.1 cc/10 g body weight). Mice were left undisturbed in their home cages after neural tagging to ensure minimal non-specific recombination.

Open field test—Mice were acclimated to the room for 30 minutes before the open field experiment. After tethering implanted ferrules to patch cords, each subject was placed in the middle of a 42 cm (length) × 42 cm (width) × 30 cm (height) square open field arena and

movements were detected via 15 infrared beams/detectors on each side. Each test consisted of two alternating and counterbalanced light-on (473 nm, 7–9 mW delivered at 20 Hz, 10 ms pulses) and light-off periods lasting 5 minutes each, for a total of 20 min. Infrared beam breaks and locomotor parameters were quantified using Fusion v5.6 SuperFlex software and reported as averages of the two light-on or light-off epochs. For analysis of morphine-induced locomotion, WT mice were injected with morphine (10 mg/kg) or saline and immediately placed in the open field for 60 min.

Conditioned place preference and aversion—The experimental apparatus was a rectangular arena with three compartments (paired, center, and unpaired) and featured distinct visual and tactile cues. SST-FlpO mice were subjected to fear conditioning or morphine administration followed by either vehicle or 4-hydroxytamoxifen injection. Following viral expression, CPP and CPA analysis was performed. On day 1 (pre-test), mice were placed in the arena and allowed to freely explore for 20 minutes. Data were scored to determine the amount of time mice spent in each compartment. On days 2–4 (conditioning), dividers were inserted into the arena to restrict mice to a single compartment. For conditioned place preference, morphine-tagged mice were placed in the compartment opposite the one for which they expressed an initial preference during the pre-test. For conditioned place avoidance, fear-tagged mice were placed in the compartment for which they expressed an initial preference during the pre-test. Mice that did not display an overt initial preference were placed into a randomly chosen compartment. Animals were then subjected to optogenetic stimulation for 40 minutes (473 nm, 7–9 mW delivered at 20 Hz, 10 ms pulses; alternating 1 min on and 1 min off). To control for independent effects of context exposure, each subject was also placed in the opposite compartment for 40 min without laser stimulation. Conditioning sessions with and without photostimulation were separated by 5–6 hours and the order was counterbalanced. On day 5 (post-test), mice were placed back into the arena with the dividers removed and allowed once again to freely explore all compartments for 20 min. Mice were tethered to patch cords throughout all test and conditioning sessions. Videos were acquired and automatically scored using Ethovision. CPP/CPA score was calculated using the following formula: (percent time in preferred side in the post-test – percent time in the same side in the pre-test) / percent time in the same side during the pre-test.

Immunohistochemistry—For cFos analysis of fear memory retrieval, mice were submitted to 4 CS presentations in a neutral context. For cFos analysis following optogenetic stimulation of tagged SST-INs (Figure 7), animals received 6 photostimulation epochs (473nm, 20 Hz, 10 ms pulses) lasting 20 s. 90 min following CS presentation or photoexcitation mice were deeply anesthetized and transcardially perfused with phosphate buffered saline (PBS) followed by 4% paraformaldehyde (PFA; pH 7.45). Brains were removed and fixed overnight in 4% PFA before sectioning at 50 μm thickness on the coronal plane using a Leica VT1000S vibratome. Immunohistochemistry was performed on floating sections with an anti-cFos primary antibody (1:1000; Millipore ABE457). Slices were blocked at room temperature for one hour by using 2% normal goat serum in PBS + 0.3% Tween-20 before overnight incubation in primary antibody at 4°C. Slices were then incubated for 2 hours at room temperature in a secondary goat anti-rabbit antibody

conjugated to Alexa 647 (1:500; Jackson Immunoresearch 111–605-045) in PBS containing 2% normal goat serum and 0.3% Tween-20. Staining against somatostatin was performed using a rabbit anti-somatostatin-14 antibody (1:1000; Peninsula Laboratories Cat. # T-4103) in PBS containing 5% BSA and 0.25% Triton X-100 following blocking in the same solution. Staining against parvalbumin and vasoactive intestinal peptide was performed using a mouse anti-parvalbumin (1:1000; Millipore Cat. # MAB1572) and a rabbit anti-VIP antibody (1:500; Immunostar Cat. # 20077), respectively, in PBS containing 5% normal goat serum and 0.3% Tween-20 after blocking in the same solution. Secondary antibody labeling was done using goat anti-rabbit conjugated to Alexa 647 (1:500; Jackson Immunoresearch 111–605-045) for tissue previously labeled for SST and VIP. Secondary antibody labeling was done using goat anti-mouse conjugated to Alexa 647 (1:500; Jackson Immunoresearch 115–605-003) for tissue previously labeled for PV. Slices were mounted with Prolong Antifade Gold mounting medium with DAPI (Life Technologies, Grand Island, NY, USA) and imaged using an inverted Zeiss 780 confocal microscope running Zen Black software (Carl Zeiss Microscopy, Jena, Germany). Cell counts were performed using the Cell Counter plug-in in ImageJ (NIH, Bethesda, MD, USA) by an experimenter blind to treatment condition.

Slice electrophysiology—Mice were deeply anesthetized with isoflurane and brains were extracted and submerged in ice cold (−2° to −4°C) carbogen-bubbled (95% oxygen, 5% CO2) sucrose cutting solution containing (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 $CaCl_2$. Coronal sections were obtained from mPFC at 300 μm thickness and recovered for 45 min at 35°C in carbogen-bubbled artificial cerebrospinal fluid (ACSF) containing (in mM): 119 NaCl, 26.2 $NaHCO_3$, 11 glucose, 2.5 KCl, 1 NaH_2PO_4, 2 $MgCl_2$, and 2 $CaCl_2$. Slices were subsequently maintained and recordings were performed at room temperature. Whole-cell electrodes (2–5 MΩ) were fabricated from borosilicate glass and filled with internal solution (pH 7.25; 295 mOsmol) containing (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 for voltage-clamp experiments. An internal solution (pH 7.25; 295–300 mOsmol) containing (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 was used for current-clamp experiments. Experiments were conducted on an upright microscope equipped with DIC optics and a light-emitting diode (LED) -coupled 40X objective. Tagged neurons were targeted based on eYFP and tdTomato fluorescence. In addition, interneurons were differentiated from PNs on the basis of morphology, capacitance and input resistance. Recordings of EPSCs and IPSCs were obtained in standard ACSF at −60 mV and 0 mV, respectively. Evoked EPSCs were produced with a bipolar stimulating electrode placed in layer 2 of prelimbic cortex. For light-evoked IPSCs, cells were clamped at 0 mV in standard ACSF with the addition of 1 μM TTX (Abcam) and 100 μM 4-aminopyrimidine (Abcam) and currents were elicited with a TTL-pulsed objective-coupled LED (460 nm, 20 mW/mm2, 1 ms pulse, Prizmatix). Data were acquired at 10 kHz and low-pass filtered at 10 kHz for spontaneous and 3kHz for evoked responses using Multiclamp 700B (Molecular Devices, San Jose, CA, USA) and pClamp 10 software (Molecular Devices). For analysis of intrinsic firing properties, recordings were performed in standard ACSF with addition of CNQX (10 μM)

and picrotoxin (100 μM). Current injections (500 ms duration) were performed at −20 pA to +90 pA at 10 pA increments. Rheobase was defined as the minimum current level required to elicit a single action potential. Cells exhibiting unacceptable health (>100 pA holding current), access resistance changes of >20% or unstable synaptic responses were excluded from analysis. All analyses were performed in ClampFit 10 (Molecular Devices) and MiniAnalysis (Synaptosoft, Fort Lee, NJ) by an experimenter blind to group allocation, cell type, and sex.

Quantification and statistical analysis

As preconditions for parametric statistical analysis, we tested the assumptions of normality and homogeneity of variance using the Shapiro-Wilk and Levene's tests, respectively. When these assumptions were not met ($p < 0.05$) non-parametric alternatives were used. Due to the lack of a non-parametric alternative to the 2-way ANOVA, this precluded testing for interactions in some experiments where two factors were present. Following a significant omnibus test, pairwise comparisons among all the groups were used to test for significant differences. All statistical tests, values, and significance levels are specified in the associated figure legends. Box plots depict median (line), mean (open square), quartiles (box), and 10–90% range (whiskers). Individual data points (open circles) are overlaid on box plots. Sample sizes were estimated based on prior electrophysiological and behavioral studies from our laboratory. Statistical analysis and graph construction were done in Graphpad Prism (San Diego, CA) and OriginPRO (OriginLab, Northampton, MA).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Anosha Khawaja for animal care, Dr. Haruhiko Bito for sharing the E-SARE-ERCreER plasmid, and members of the Clem laboratory for helpful discussions. We also thank Dr. Marine Salery for advice regarding CPP experiments, and Dr. Abigale Lade for her immeasurable support. These experiments were supported by funds from the National Institute of Mental Health (NIMH) grants R01 MH116145 and R01 MH124880 to R.L.C., in addition to F32 MH115688 and K99 MH122228 to K.A.C.

References

Alberini CM (2009). Transcription factors in long-term memory and synaptic plasticity. Physiol Rev 89, 121–145. 10.1152/physrev.00017.2008. [PubMed: 19126756]

Artinian J, and Lacaille JC (2018). Disinhibition in learning and memory circuits: New vistas for somatostatin interneurons and long-term synaptic plasticity. Brain Research Bulletin 141, 20–26. 10.1016/j.brainresbull.2017.11.012. [PubMed: 29174732]

Bagur S, Lefort JM, Lacroix MM, de Lavilleon G, Herry C, Chouvaeff M, Billand C, Geoffroy H, and Benchenane K (2021). Breathing-driven prefrontal oscillations regulate maintenance of conditioned-fear evoked freezing independently of initiation. Nat Commun 12, 2605. 10.1038/s41467-021-22798-6. [PubMed: 33972521]

Beas BS, Wright BJ, Skirzewski M, Leng Y, Hyun JH, Koita O, Ringelberg N, Kwon HB, Buonanno A, and Penzo MA (2018). The locus coeruleus drives disinhibition in the midline thalamus via a dopaminergic mechanism. Nat Neurosci 21, 963–973. 10.1038/s41593-018-0167-4. [PubMed: 29915192]

Bocchio M, Nabavi S, and Capogna M (2017). Synaptic Plasticity, Engrams, and Network Oscillations in Amygdala Circuits for Storage and Retrieval of Emotional Memories. Neuron 94, 731–743. 10.1016/j.neuron.2017.03.022. [PubMed: 28521127]

Burgos-Robles A, Vidal-Gonzalez I, and Quirk GJ (2009). Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. J Neurosci 29, 8474–8482. 10.1523/JNEUROSCI.0378-09.2009. [PubMed: 19571138]

Cai DJ, Aharoni D, Shuman T, Shobe J, Biane J, Song W, Wei B, Veshkini M, La-Vu M, Lou J, et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. Nature 534, 115–118. [PubMed: 27251287]

Clem RL, and Schiller D (2016). New Learning and Unlearning: Strangers or Accomplices in Threat Memory Attenuation? Trends Neurosci 39, 340–351. 10.1016/j.tins.2016.03.003. [PubMed: 27079843]

Corcoran KA, and Quirk GJ (2007). Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. J Neurosci 27, 840–844. 10.1523/JNEUROSCI.5327-06.2007. [PubMed: 17251424]

Courtin J, Chaudun F, Rozeske RR, Karalis N, Gonzalez-Campo C, Wurtz H, Abdi A, Baufreton J, Bienvenu TC, and Herry C (2014). Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. Nature 505, 92–96. 10.1038/nature12755. [PubMed: 24256726]

Cummings KA, and Clem RL (2020). Prefrontal somatostatin interneurons encode fear memory. Nat Neurosci 23, 61–74. 10.1038/s41593-019-0552-7. [PubMed: 31844314]

Dejean C, Courtin J, Karalis N, Chaudun F, Wurtz H, Bienvenu TC, and Herry C (2016). Prefrontal neuronal assemblies temporally control fear behaviour. Nature 535, 420–424. 10.1038/nature18630. [PubMed: 27409809]

DeNardo LA, Liu CD, Allen WE, Adams EL, Friedmann D, Fu L, Guenthner CJ, Tessier-Lavigne M, and Luo L (2019). Temporal evolution of cortical ensembles promoting remote memory retrieval. Nat Neurosci 22, 460–469. 10.1038/s41593-018-0318-7. [PubMed: 30692687]

Do-Monte FH, Quinones-Laracuente K, and Quirk GJ (2015). A temporal shift in the circuits mediating retrieval of fear memory. Nature 519, 460–463. 10.1038/nature14030. [PubMed: 25600268]

Funk CM, Peelman K, Bellesi M, Marshall W, Cirelli C, and Tononi G (2017). Role of Somatostatin-Positive Cortical Interneurons in the Generation of Sleep Slow Waves. J Neurosci 37, 9132–9148. 10.1523/JNEUROSCI.1303-17.2017. [PubMed: 28821651]

Garcia-Junco-Clemente P, Ikrar T, Tring E, Xu X, Ringach DL, and Trachtenberg JT (2017). An inhibitory pull-push circuit in frontal cortex. Nat Neurosci 20, 389–392. 10.1038/nn.4483. [PubMed: 28114295]

Han JH, Kushner SA, Yiu AP, Hsiang HL, Buch T, Waisman A, Bontempi B, Neve RL, Frankland PW, and Josselyn SA (2009). Selective erasure of a fear memory. Science 323, 1492–1496. 10.1126/science.1164139. [PubMed: 19286560]

Headley DB, and Pare D (2017). Common oscillatory mechanisms across multiple memory systems. NPJ Sci Learn 2. 10.1038/s41539-016-0001-2.

Herry C, and Johansen JP (2014). Encoding of fear learning and memory in distributed neuronal circuits. Nat Neurosci 17, 1644–1654. 10.1038/nn.3869. [PubMed: 25413091]

Hyman SE, Malenka RC, and Nestler EJ (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. Annu Rev Neurosci 29, 565–598. 10.1146/annurev.neuro.29.051605.113009. [PubMed: 16776597]

Jackson J, Karnani MM, Zemelman BV, Burdakov D, and Lee AK (2018). Inhibitory Control of Prefrontal Cortex by the Claustrum. Neuron 99, 1029–1039 e1024. 10.1016/j.neuron.2018.07.031. [PubMed: 30122374]

Jiang C, Wang X, Le Q, Liu P, Liu C, Wang Z, He G, Zheng P, Wang F, and Ma L (2021). Morphine coordinates SST and PV interneurons in the prelimbic cortex to disinhibit pyramidal neurons and enhance reward. Mol Psychiatry 26, 1178–1193. 10.1038/s41380-019-0480-7. [PubMed: 31413370]

Johnson PL, and Shekhar A (2006). Panic-prone state induced in rats with GABA dysfunction in the dorsomedial hypothalamus is mediated by NMDA receptors. J Neurosci 26, 7093–7104. 10.1523/JNEUROSCI.0408-06.2006. [PubMed: 16807338]

Josselyn SA, and Tonegawa S (2020). Memory engrams: Recalling the past and imagining the future. Science 367, eaaw4325–eaaw4325. [PubMed: 31896692]

Karalis N, Dejean C, Chaudun F, Khoder S, Rozeske RR, Wurtz H, Bagur S, Benchenane K, Sirota A, Courtin J, and Herry C (2016). 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. Nat Neurosci 19, 605–612. 10.1038/nn.4251. [PubMed: 26878674]

Kawashima T, Kitamura K, Suzuki K, Nonaka M, Kamijo S, Takemoto-Kimura S, Kano M, Okuno H, Ohki K, and Bito H (2013). Functional labeling of neurons and their projections using the synthetic activity-dependent promoter E-SARE. Nat Methods 10, 889–895. 10.1038/nmeth.2559. [PubMed: 23852453]

Kawashima T, Okuno H, and Bito H (2014). A new era for functional labeling of neurons: activity-dependent promoters have come of age. Front Neural Circuits 8, 37. 10.3389/fncir.2014.00037. [PubMed: 24795570]

Kim WB, and Cho JH (2017). Encoding of Discriminative Fear Memory by Input-Specific LTP in the Amygdala. Neuron 95, 1129–1146 e1125. 10.1016/j.neuron.2017.08.004. [PubMed: 28823727]

Krabbe S, Paradiso E, d'Aquin S, Bitterman Y, Courtin J, Xu C, Yonehara K, Markovic M, Muller C, Eichlisberger T, et al. (2019). Adaptive disinhibitory gating by VIP interneurons permits associative learning. Nat Neurosci 22, 1834–1843. 10.1038/s41593-019-0508-y. [PubMed: 31636447]

Lacagnina AF, Brockway ET, Crovetti CR, Shue F, McCarty MJ, Sattler KP, Lim SC, Santos SL, Denny CA, and Drew MR (2019). Distinct hippocampal engrams control extinction and relapse of fear memory. Nature Neuroscience 22, 753–761. 10.1038/s41593-019-0361-z. [PubMed: 30936555]

Le Merrer J, Becker JA, Befort K, and Kieffer BL (2009). Reward processing by the opioid system in the brain. Physiol Rev 89, 1379–1412. 10.1152/physrev.00005.2009. [PubMed: 19789384]

Lee KH, Chai CY, Wayner MJ, Chung PM, and Hsu CH (1977). Effects of neuroleptics on morphine-induced tail erection in mice. Pharmacol Biochem Behav 7, 153–157. 10.1016/0091-3057(77)90200-3. [PubMed: 21420]

Letzkus JJ, Wolff SB, and Luthi A (2015). Disinhibition, a Circuit Mechanism for Associative Learning and Memory. Neuron 88, 264–276. 10.1016/j.neuron.2015.09.024. [PubMed: 26494276]

Li K, Nakajima M, Ibanez-Tallon I, and Heintz N (2016). A Cortical Circuit for Sexually Dimorphic Oxytocin-Dependent Anxiety Behaviors. Cell 167, 60–72 e11. 10.1016/j.cell.2016.08.067. [PubMed: 27641503]

Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, and Tonegawa S (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484, 381–385. [PubMed: 22441246]

Lucas EK, and Clem RL (2017). GABAergic interneurons_ The orchestra or the conductor in fear learning and memory? Brain Research Bulletin, 0–1.

Lucas EK, and Clem RL (2018). GABAergic interneurons: The orchestra or the conductor in fear learning and memory? Brain Res Bull 141, 13–19. 10.1016/j.brainresbull.2017.11.016. [PubMed: 29197563]

Malenka RC, and Bear MF (2004). LTP and LTD: an embarrassment of riches. Neuron 44, 5–21. 10.1016/j.neuron.2004.09.012. [PubMed: 15450156]

Nakajima M, Gorlich A, and Heintz N (2014). Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. Cell 159, 295–305. 10.1016/j.cell.2014.09.020. [PubMed: 25303526]

Pi H-J, Hangya B, Kvitsiani D, Sanders JI, Huang ZJ, and Kepecs A (2013). Cortical interneurons that specialize in disinhibitory control. Nature 503, 521–524. [PubMed: 24097352]

Rashid AJ, Yan C, Mercaldo V, Hsiang HL, Park S, Cole CJ, De Cristofaro A, Yu J, Ramakrishnan C, Lee SY, et al. (2016). Competition between engrams influences fear memory formation and recall. Science 353, 383–387. 10.1126/science.aaf0594. [PubMed: 27463673]

Reijmers LG, Perkins BL, Matsuo N, and Mayford M (2007). Localization of a stable neural correlate of associative memory. Science 317, 1230–1233. 10.1126/science.1143839. [PubMed: 17761885]

Ressler RL, and Maren S (2019). Synaptic encoding of fear memories in the amygdala. Curr Opin Neurobiol 54, 54–59. 10.1016/j.conb.2018.08.012. [PubMed: 30216780]

Rosen LG, Sun N, Rushlow W, and Laviolette SR (2015). Molecular and neuronal plasticity mechanisms in the amygdala-prefrontal cortical circuit: implications for opiate addiction memory formation. Front Neurosci 9, 399. 10.3389/fnins.2015.00399. [PubMed: 26594137]

Rudy B, Fishell G, Lee S, and Hjerling-Leffler J (2011). Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol 71, 45–61. 10.1002/dneu.20853. [PubMed: 21154909]

Ryan TJ, Roy DS, Pignatelli M, Arons A, and Tonegawa S (2015). Memory. Engram cells retain memory under retrograde amnesia. Science 348, 1007–1013. 10.1126/science.aaa5542. [PubMed: 26023136]

Sotres-Bayon F, and Quirk GJ (2010). Prefrontal control of fear: more than just extinction. Curr Opin Neurobiol 20, 231–235. 10.1016/j.conb.2010.02.005. [PubMed: 20303254]

Stamatakis AM, and Stuber GD (2012). Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. Nat Neurosci 15, 1105–1107. 10.1038/nn.3145. [PubMed: 22729176]

Tayler KK, Tanaka KZ, Reijmers LG, and Wiltgen BJ (2013). Reactivation of neural ensembles during the retrieval of recent and remote memory. Current Biology 23, 99–106. 10.1016/j.cub.2012.11.019. [PubMed: 23246402]

Tonegawa S, Liu X, Ramirez S, and Redondo R (2015). Memory Engram Cells Have Come of Age. Neuron 87, 918–931. 10.1016/j.neuron.2015.08.002. [PubMed: 26335640]

Tovote P, Esposito MS, Botta P, Chaudun F, Fadok JP, Markovic M, Wolff SBE, Ramakrishnan C, Fenno L, Deisseroth K, et al. (2016). Midbrain circuits for defensive behaviour. Nature, 1–17.

Wolff SB, Grundemann J, Tovote P, Krabbe S, Jacobson GA, Muller C, Herry C, Ehrlich I, Friedrich RW, Letzkus JJ, and Luthi A (2014). Amygdala interneuron subtypes control fear learning through disinhibition. Nature 509, 453–458. 10.1038/nature13258. [PubMed: 24814341]

Xu H, Jeong HY, Tremblay R, and Rudy B (2013). Neocortical somatostatin-expressing GABAergic interneurons disinhibit the thalamorecipient layer 4. Neuron 77, 155–167. 10.1016/j.neuron.2012.11.004. [PubMed: 23312523]

Xu H, Liu L, Tian Y, Wang J, Li J, Zheng J, Zhao H, He M, Xu TL, Duan S, and Xu H (2019). A Disinhibitory Microcircuit Mediates Conditioned Social Fear in the Prefrontal Cortex. Neuron 102, 668–682 e665. 10.1016/j.neuron.2019.02.026. [PubMed: 30898376]

Yamamuro K, Bicks LK, Leventhal MB, Kato D, Im S, Flanigan ME, Garkun Y, Norman KJ, Caro K, Sadahiro M, et al. (2020). A prefrontal-paraventricular thalamus circuit requires juvenile social experience to regulate adult sociability in mice. Nat Neurosci 23, 1240–1252. 10.1038/s41593-020-0695-6. [PubMed: 32868932]

Yavorska I, and Wehr M (2016). Somatostatin-Expressing Inhibitory Interneurons in Cortical Circuits. Front Neural Circuits 10, 76. 10.3389/fncir.2016.00076. [PubMed: 27746722]

Ye L, Allen WE, Thompson KR, Tian Q, Hsueh B, Ramakrishnan C, Wang AC, Jennings JH, Adhikari A, Halpern CH, et al. (2016). Wiring and Molecular Features of Prefrontal Ensembles Representing Distinct Experiences. Cell 165, 1776–1788. 10.1016/j.cell.2016.05.010. [PubMed: 27238022]

Highlights

- A discrete subset of prefrontal SST-INs participates in fear memory encoding.
- Memory-encoding SST-INs have distinct circuit properties.
- Morphine activates an orthogonal SST-IN population that opposes cued fear.
- Fear- and morphine-related SST-INs modulate valence-specific brain networks.

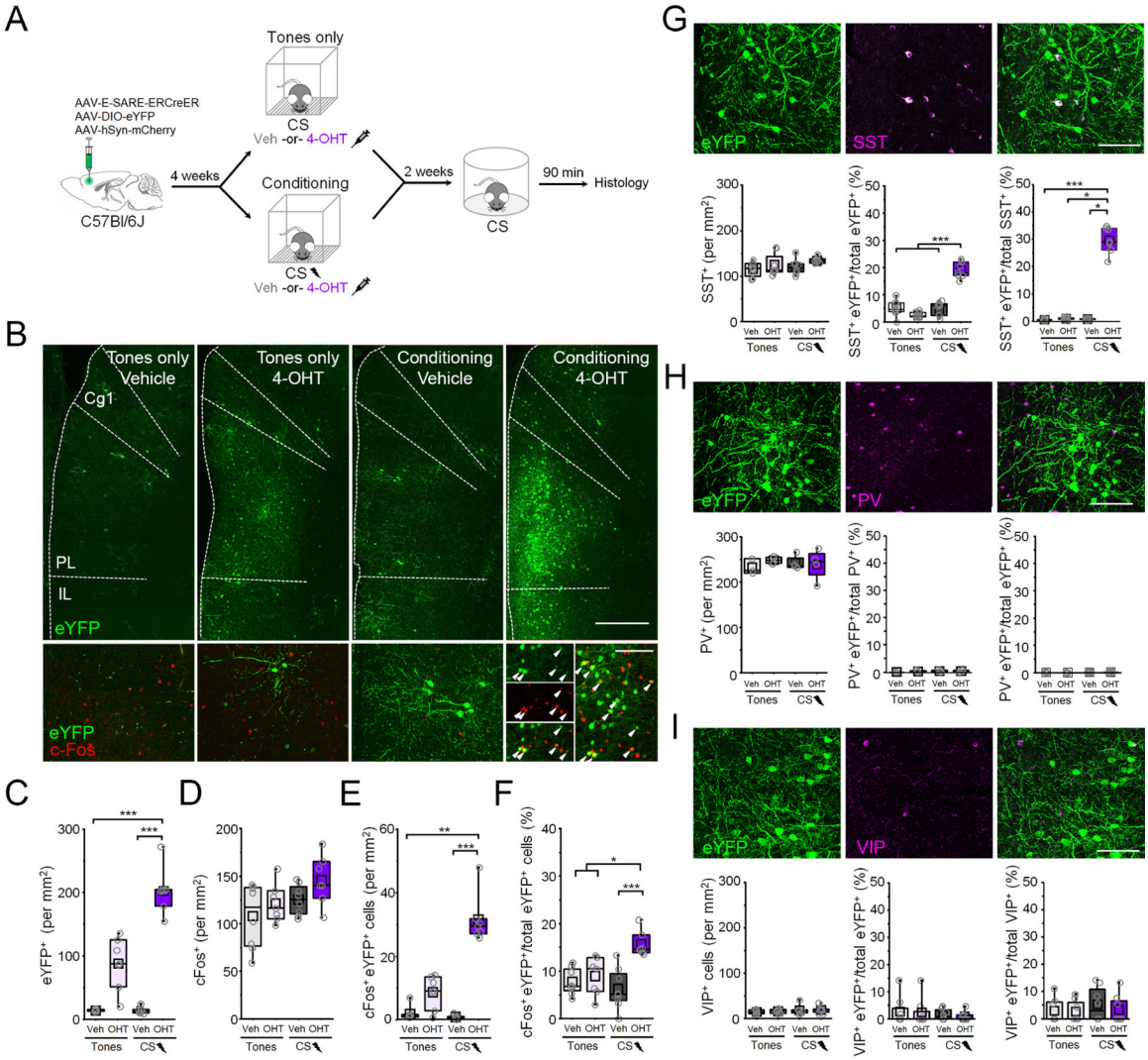


Figure 1. Fear conditioning recruits a cell population that is reactivated upon retrieval and is partially comprised of somatostatin-expressing interneurons.

(A) Timeline for cellular tagging during fear conditioning followed by cFos analysis of retrieval-dependent reactivation. **(B)** Top: representative eYFP tagging. Scale bar = 500 μm. Bottom: cFos induction following the retrieval test. White arrowheads denote cFos+ eYFP neurons. Scale bar = 100 μm. Cg1 = cingulate area 1. PL = prelimbic cortex. IL = infralimbic cortex. **(C-F)** Boxplot comparison between tones only vehicle (n = 8 mice), tones only 4-OHT (n = 7 mice), conditioned vehicle (n = 8 mice), and conditioned 4-OHT (n = 7 mice) groups of **(C)** number of eYFP+ cells: $\chi^2 = 23.7$ (3), $p = 2.89 \times 10^{-5}$, Kruskal-Wallis ANOVA; **(D)** number of cFos+ cells: $\chi^2 = 5.75$ (3), $p = 0.12$, Kruskal-Wallis ANOVA; **(E)** number of cFos+/eYFP+ double positive cells: $\chi^2 = 20.9$ (3), $p = 1.09 \times 10^{-4}$, Kruskal-Wallis ANOVA; and **(F)** number of cFos+/eYFP+ double positive cells normalized to the total number of eYFP+ cells: $\chi^2 = 16.7$ (3), $p = 8.05 \times 10^{-4}$, Kruskal-Wallis ANOVA. **(G)** SST labeling following CS-US pairing and 4-OHT injection. Scale = 50 μm. Bottom, boxplot comparison of labeling in tones only vehicle (n = 8 mice), tones only 4-OHT (n = 8 mice), conditioning vehicle (n = 8 mice), and conditioning 4-OHT (n = 7 mice). SST+

cells: χ^2 = 6.93 (3), p = 0.074, Kruskal-Wallis ANOVA. SST+ eYFP cells normalized to the total number of eYFP cells: $F_{(1,27)}$ = 87.9, p = 5.55 × 10^{-10}, interaction between training and treatment, 2-way ANOVA. SST+ eYFP cells normalized to the total number of SST+ cells per group: χ^2 = 17.04 (3), p = 6.93 × 10^{-4}, Kruskal-Wallis ANOVA. **(H)** PV labeling following CS-US pairing and 4-OHT injection. Scale = 50 μm. Bottom, boxplot comparison of labeling in tones only vehicle (n = 3), tones only 4-OHT (n = 4), conditioning vehicle (n = 4), and conditioning 4-OHT (n = 4). PV+ cells: $F_{(1,40)}$ = 0.73, p = 0.41, 2-way ANOVA. PV+ eYFP cells normalized to the total number of eYFP+ cells per group: χ^2 = 0.88 (3), p = 0.83, Kruskal-Wallis ANOVA. PV+ eYFP cells normalized to the total number of PV+ cells per group: χ^2 = 1.31 (3), p = 0.73, Kruskal-Wallis ANOVA. **(I)** VIP labeling following CS-US pairing and 4-OHT injection. Scale = 50 μm. Bottom, boxplot comparison between tones only vehicle (n = 8), tones only 4-OHT (n = 8), conditioning vehicle (n = 8), and conditioning 4-OHT (n = 8). VIP+ cells: χ^2 = 1.84 (3), p = 0.606, Kruskal-Wallis ANOVA. VIP+ eYFP cells normalized to the total number of eYFP+ cells per group: χ^2 = 0.359 (3), p = 0.949, Kruskal-Wallis ANOVA. VIP+ eYFP cells normalized to the total number of VIP+ cells per group: χ^2 = 1.02 (3), p = 0.796, Kruskal-Wallis ANOVA. * p < 0.05, ** p < 0.01, *** p < 0.001 by Tukey's post-hoc test (**G**) or Dunn's post-hoc test (**C, E, F, G**). See also Figures S1-S3.

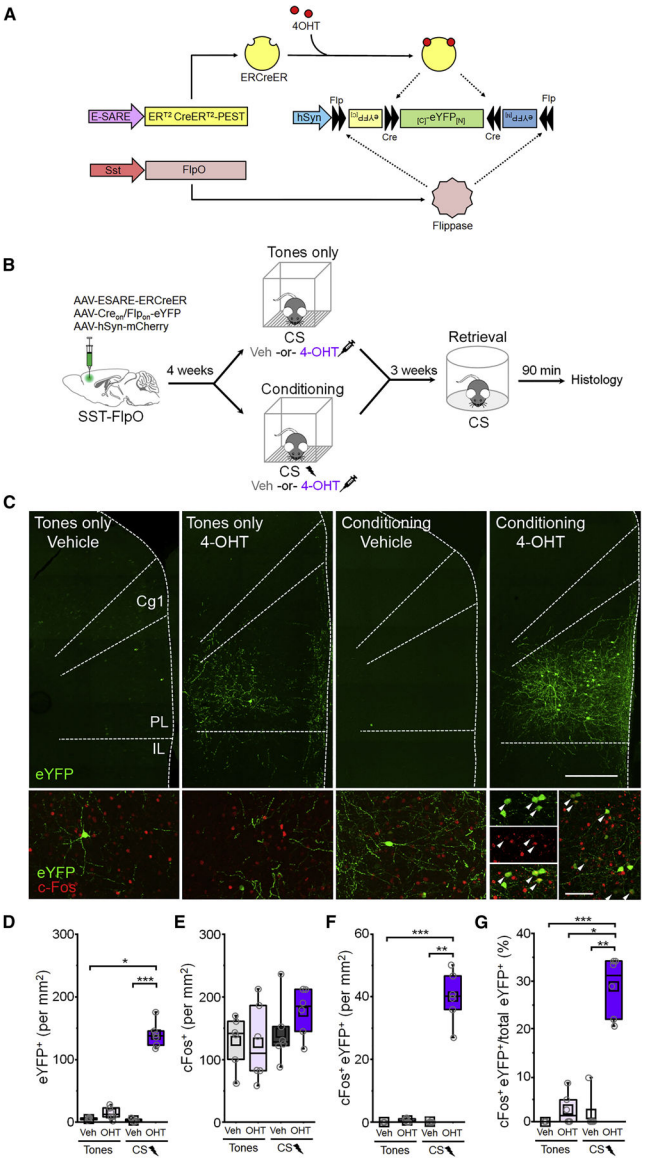


Figure 2: Fear learning-activated somatostatin-interneurons are preferentially reactivated upon memory retrieval.

(A) Intersectional genetic targeting. **(B)** Timeline for cellular tagging followed by cFos analysis of retrieval-dependent reactivation. **(C)** Top: representative eYFP tagging. Scale bar = 500 μm. Bottom: cFos induction following retrieval. White arrowheads denote cFos+ eYFP neurons. Scale bar = 100 μm. Cg1 = cingulate area 1. PL = prelimbic cortex. IL = infralimbic cortex. **(D-G)** Boxplot comparison between tones only vehicle (n = 6 mice), tones only 4-OHT (n = 6 mice), conditioning vehicle (n = 6 mice), and conditioning 4-OHT (n = 6 mice) groups of **(D)** number of eYFP+ cells: $\chi^2 = 17.6$ (3), $p = 5.20 \times 10^{-4}$, Kruskal-Wallis ANOVA; **(E)** number of cFos+ cells: number of cFos+ cells: $F_{(1,20)} = 0.82$, p = 0.38, 2-way ANOVA. **(F)** number of cFos+/eYFP+ double positive cells: $\chi^2 = 18.0$ (3), $p = 4.46 \times 10^{-4}$, Kruskal-Wallis ANOVA; and **(G)** number of cFos+/eYFP+ double positive cells normalized to the total number of eYFP+ cells in each group: $\chi^2 = 17.5$ (3), p = 5.49

$\times 10^{-4}$, Kruskal-Wallis ANOVA. $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Dunn's post-hoc test **(D,F,G)**. See also Figure S4.

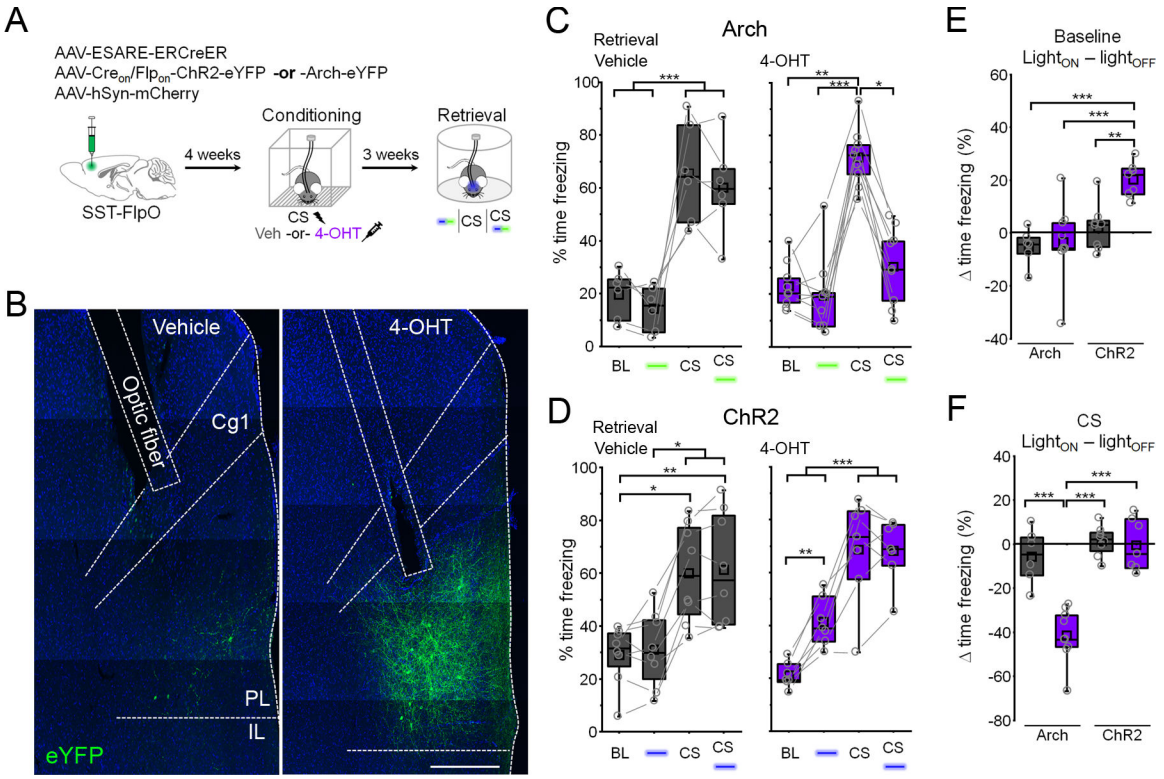


Figure 3: Fear learning-activated somatostatin interneurons control cue-elicited freezing.

(A) Timeline for cellular tagging and optogenetic analysis of fear behavior. **(B)** Representative ChR2 expression and optic fiber placement. Scale = 500 μm. Cg1 = cingulate area 1. PL = prelimbic cortex. IL = infralimbic cortex. **(C)** Boxplots depict modulation of freezing by light (532 nm, constant, 20 s epochs) and CS trials in vehicle (gray) and 4-OHT (purple) injected mice expressing Arch. Vehicle: $F_{(3,15)} = 22.1$, $p = 9.18 \times 10^{-6}$, 1-way repeated measures ANOVA, n = 6 mice. 4-OHT: $\chi^2 = 17.4$ (3), $p = 5.85 \times 10^{-4}$, Friedman ANOVA, n = 9 mice. (**D**) Boxplots depict modulation of freezing by light (473 nm, 10 ms pulses, 20 Hz, 20 s epochs) and CS trials in vehicle (gray) and 4-OHT (purple) injected mice expressing ChR2. Vehicle: $\chi^2 = 19.5$ (3), $p = 2.15 \times 10^{-4}$, Friedman ANOVA, n = 8 mice. 4-OHT: $F_{(3,18)} = 46.7$, $p = 1.07 \times 10^{-8}$, 1-way repeated measures ANOVA, n = 7 mice. **(E)** Boxplots depict change in freezing induced by photostimulation during the baseline period in (**C**) and (**D**). Effect of photostimulation (Light$_{ON}$ – light$_{OFF}$): $F_{(1,26)} = 4.73$, $p = 0.039$, 2-way ANOVA. **(F)** Boxplots depict change in freezing induced by photostimulation during the CS in (**C**) and (**D**). Effect of photostimulation (Light$_{ON}$ – light$_{OFF}$): $F_{(1,26)} = 17.5$, $p = 2.87 \times 10^{-4}$, 2-way ANOVA. ** $p < 0.01$, *** $p < 0.001$, Tukey's post-hoc test (**C**: vehicle, **D**: 4-OHT, **E**, **F**). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Dunn's post-hoc test (**C**: 4-OHT, **D**: vehicle). See also Figure S5.

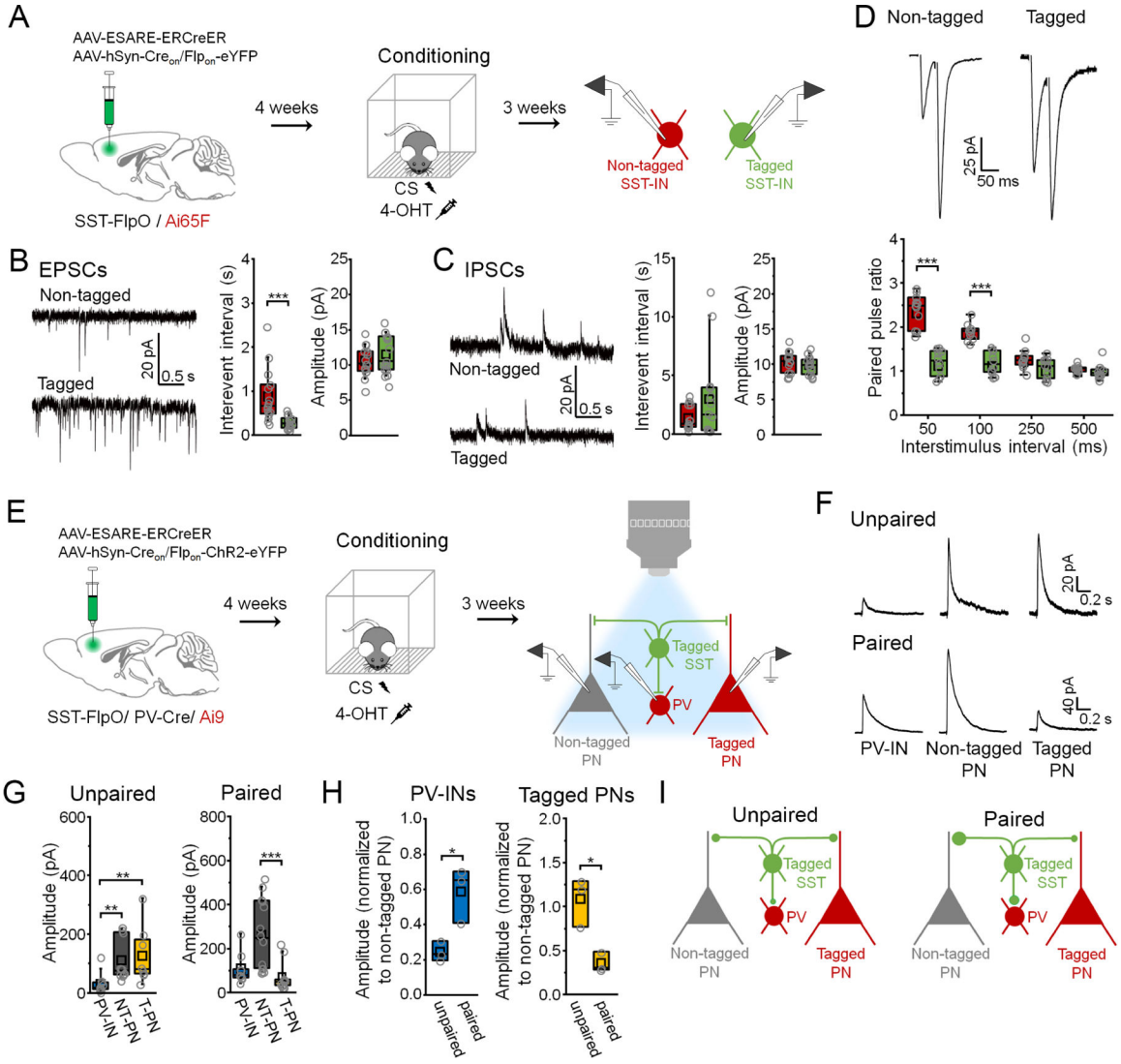


Figure 4: Input and output synaptic connections of fear learning-activated somatostatin interneurons exhibit distinct properties.

(A) Targeting of activated (eYFP+/ tdTomato+) versus nonactivated SST-INs (eYFP−/ tdTomato+) for electrophysiology following CS-US pairing. **(B)** Boxplots depict spontaneous EPSC parameters from tagged (n = 15 cells) and non-tagged SST-INs (n = 15 cells) in the same slices (n = 6 slices from 5 mice). Interevent interval: U = 205, p = 1.35×10^{-4}, Mann-Whitney U-test. Amplitude: $t_{28} = -1.03$, p = 0.31, two-sided unpaired t-test. **(C)** Boxplots depict spontaneous IPSC parameters from tagged (n = 14 cells) and non-tagged SST-INs (n = 14 cells) in the same slices (n = 6 slices from 5 mice). Interevent interval: U = 89, p = 0.696, Mann-Whitney U-test. Amplitude: $t_{26} = 0.945$, p = 0.353, Mann-Whitney U-test. **(D)** Boxplots depict EPSC amplitude from tagged (n = 10 cells) and non-tagged SST-INs (n = 11 cells) in the same slices (n = 4 slices in 3 mice) during paired pulse stimulation. Paired pulse ratio: $F_{(3,7)} = 16.6$, p = 0.0014, interaction between cell type and delay, 2-way repeated measures ANOVA. **(E)** Electrophysiological analysis of connections from tagged SST-INs onto PV-INs (tdTomato+) as well as tagged (tdTomato+)

and non-tagged PNs (tdTomato-) following CS-US pairing or unpaired conditioning. **(F)** Example IPSCs. **(G)** Boxplots depict amplitude of IPSCs resulting from photoexcitation (460 nm, 1 ms, 0.1 Hz) of tagged SST-INs in PV-INs, non-tagged PNs (NT-PNs), and tagged PNs (T-PNs) following unpaired conditioning (n = 3 slices from 3 mice: PV-INs (n = 10 cells), NT-PNs (n = 10 cells), T-PNs (n = 8 cells)) and CS-US pairing (n = 3 slices from 3 mice: PV-INs (n = 9 cells), NT-PNs (n = 14 cells), T-PNs (n = 12 cells)). Unpaired conditioning: $\chi^2 = 13.0$ (2), p = 0.0015, Kruskal-Wallis ANOVA. CS-US pairing: $\chi^2 = 17.8$ (2), $p = 1.32 \times 10^{-4}$, Kruskal-Wallis ANOVA. **(H)** Boxplots depict amplitude of IPSCs from PV-INs and tagged PNs normalized to the median amplitude of non-tagged PNs from the same slices. PV-IN amplitude: $t_4 = 3.56$, p = 0.023, two-sided unpaired t-test. Tagged PN amplitude: $t_4 = 4.19$, p = 0.013, two-sided unpaired t-test. (**I**) Model of relative output strength of SST-INs in unpaired versus paired mice. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by Mann Whitney U-test (**B**), Tukey's post-hoc test (**D**), unpaired t-test (**H**) or Dunn's post-hoc test (**G**). See also Figure S6.

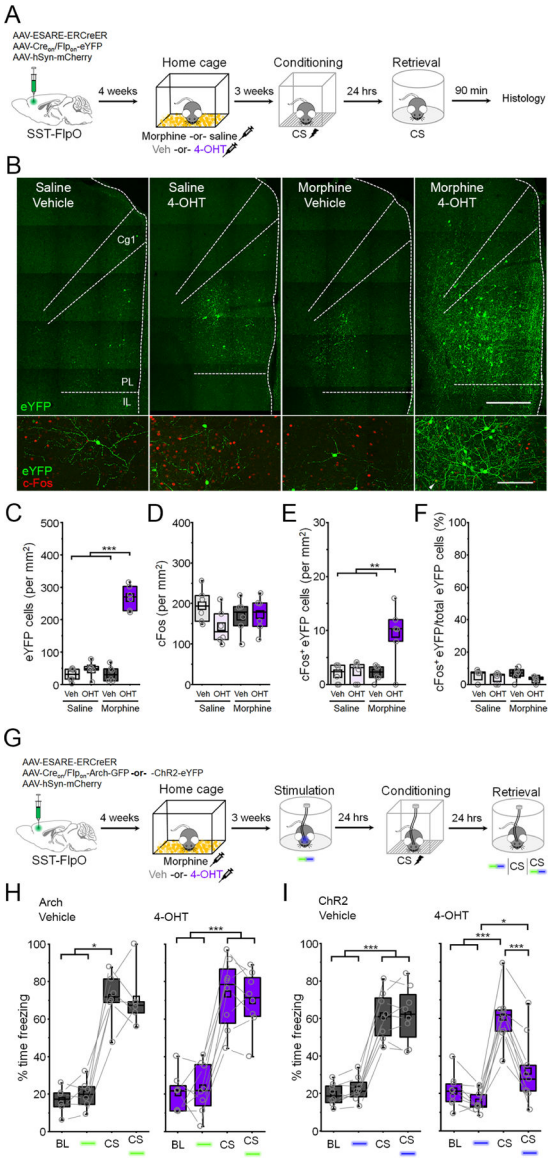


Figure 5: Morphine activates an orthogonal population of somatostatin-interneurons that opposes conditioned fear expression.

(A) Timeline for tagging of morphine-related SST-INs followed by fear conditioning and cFos analysis of retrieval-dependent reactivation. (**B**) Top: representative SST-IN tagging. Scale bar = 500 μm. Bottom: cFos induction following memory retrieval. White arrowhead denotes cFos+ eYFP neuron. Scale bar = 100 μm. Cg1 = cingulate area 1. PL = prelimbic cortex. IL = infralimbic cortex. **(C-F)** Boxplot comparison between saline vehicle (n = 6 mice), tones only 4-OHT (n = 6 mice), conditioning vehicle (n = 6 mice), and conditioning 4-OHT (n = 6 mice) groups of **(C)** number of eYFP+ cells: $F_{(1,20)} = 96.1$, $p = 4.42 \times 10^{-9}$, interaction between drug and treatment, 2-way ANOVA; **(D)** number of cFos+ cells: $F_{(1,20)} = 2.62$, $p = 0.12$, 2-way ANOVA; **(E)** number of cFos+/eYFP+ double positive cells: $F_{(1,20)} = 7.77$, $p = 0.011$, interaction between drug and treatment, 2-way ANOVA; and **(F)** number of cFos+/eYFP+ double positive cells normalized to the total number of eYFP+ cells in each group: $\chi^2 = 4.95$ (3), $p = 0.175$, Kruskal-Wallis ANOVA. **(G)** Timeline for

tagging of morphine-related SST-INs followed by optogenetic analysis of fear behavior. **(H)** Boxplots depict modulation of freezing by photoinhibition (532 nm, 20 s epochs, constant) and CS presentation during the retrieval test for vehicle (gray) and 4-OHT (purple) mice. Vehicle: $\chi^2 = 14.6$ (3), $p = 0.0022$, Friedman ANOVA. 4-OHT: $F_{(3,21)} = 66.4$, $p = 6.99 \times 10^{-11}$, 1-way repeated measures ANOVA. **(I)** Boxplots depict modulation of freezing by photoexcitation (473 nm, 20 s epochs, 20 Hz, 5 s pulses) and CS presentation during the retrieval test for vehicle (gray) and 4-OHT (purple) mice. Group x stimulus interaction: $F_{(3,21)} = 41.6$, $p = 5.11 \times 10^{-9}$, 2-way repeated measures ANOVA. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by Tukey's post-hoc test (**C, E, H, I**) or Dunn's post-hoc test (**F, H**). See also Figure S7.

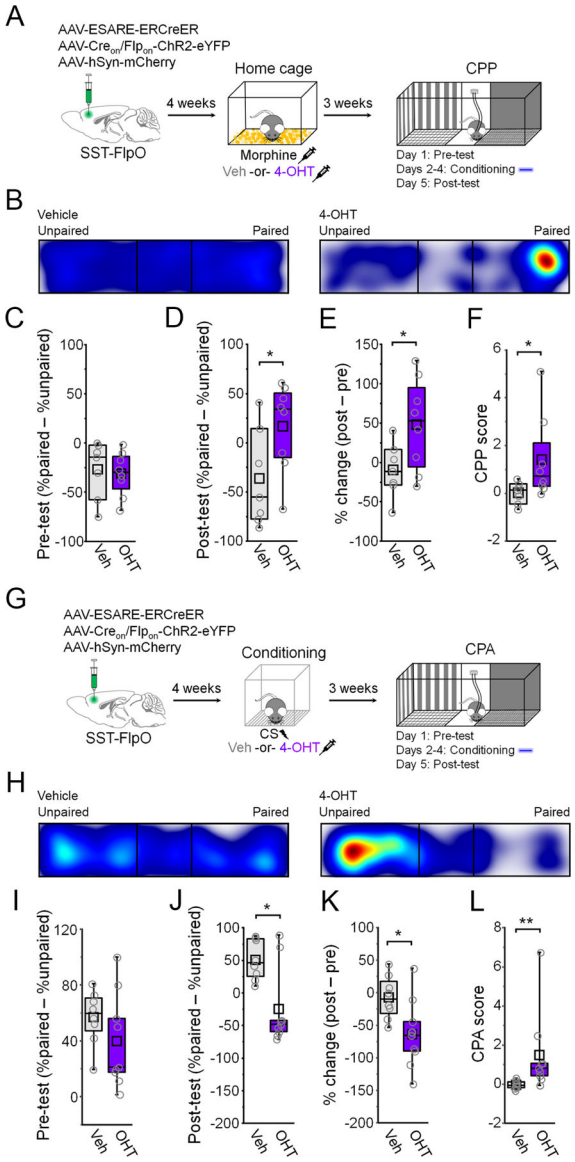


Figure 6: Morphine and fear-related SST-INs signal opposite valence.

(A) Timeline for tagging and photoexcitation of morphine-related SST-INs in a CPP assay. **(B)** Example heat maps for the post-conditioning test. **(C)** Boxplot of pre-conditioning place preference: $t_{13} = 0.32$, $p = 0.75$, two-sided unpaired t-test. **(D)** Boxplot of post-conditioning place preference: $t_{13} = 2.19$, $p = 0.047$, two-sided unpaired t-test. **(E)** Boxplot of percent change in preference: $t_{13} = 2.27$, $p = 0.041$, two-sided unpaired t-test. **(F)** CPP score: $U = 10$, $p = 0.043$, Mann-Whitney U-test. **(G)** Timeline for tagging and photoexcitation of fear-related SST-INs in a CPA assay. **(H)** Example heat maps during the post-conditioning test of place aversion. **(I)** Boxplot of pre-conditioning place preference: : $t_{15} = 1.26$, $p = 0.22$, two-sided unpaired t-test. **(J)** Boxplot of post-conditioning place preference: $U = 59$, $p = 0.031$, Mann-Whitney U-test. **(K)** Boxplot of percent change in preference: $t_{15} = 2.57$, $p = 0.021$, two-sided unpaired t-test. **(L)** CPA score: $U = 5$, $p = 0.0033$, Mann-Whitney U-test. * $p < 0.05$, ** $p < 0.01$ by unpaired t-test (**D, E, K**) or Mann-Whitney U test (**F, J, L**).

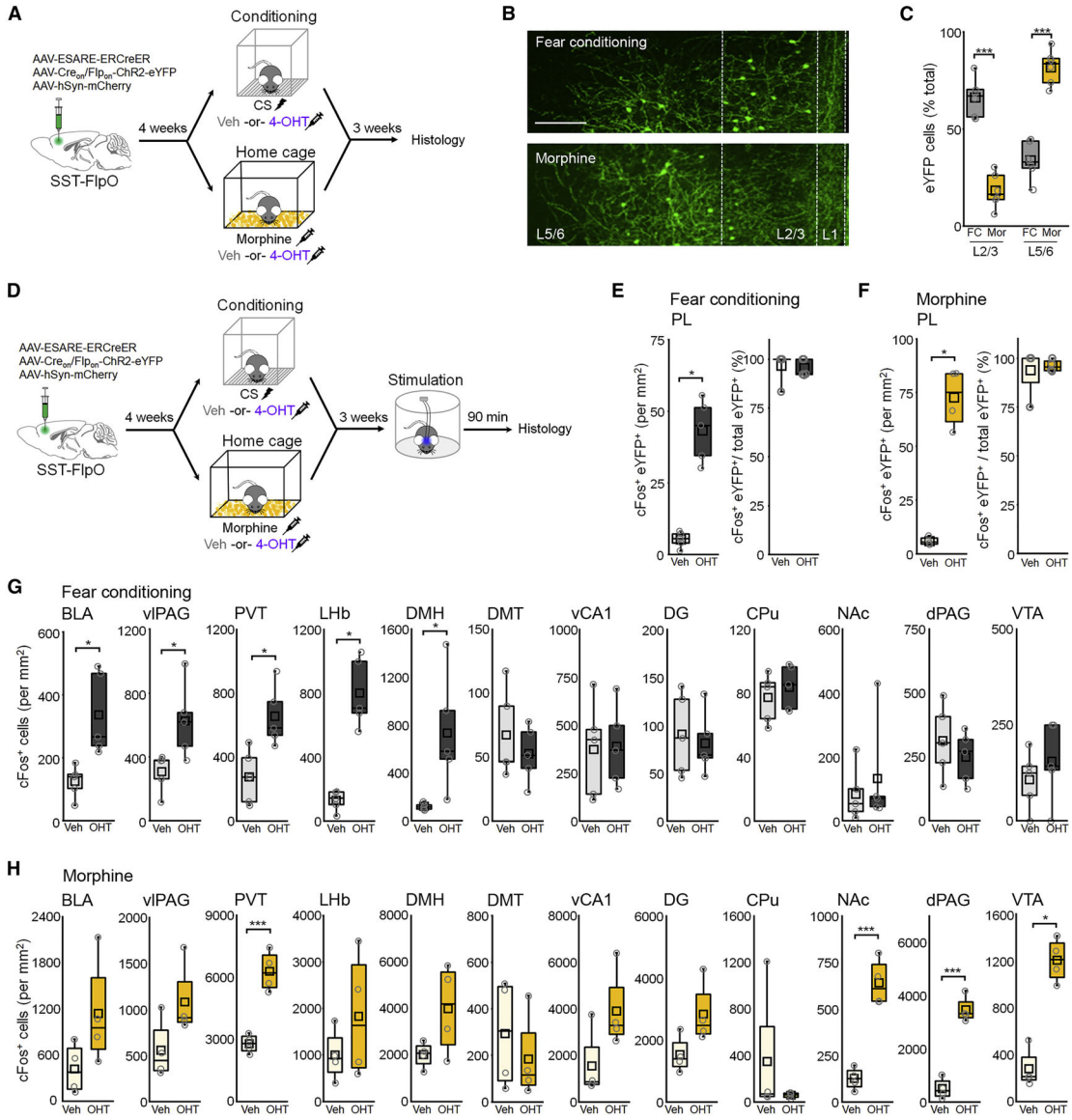


Figure 7: Fear- and morphine-related SST-INs differ in laminar distribution and modulate distinct brain networks.

(A) Timeline for cellular tagging followed by lamina-specific cell counts. **(B)** Representative eYFP tagging. Scale = 100 μm. **(C)** Boxplot of proportion of eYFP+ cells in L2/3 versus L5/6. $F_{(1,20)} = 160$, $p = 5.20 \times 10^{-11}$, 2-way ANOVA. **(D)** Timeline for cFos-based analysis of network activity following photoexcitation of fear- or morphine-related SST-INs. **(E-F)** Boxplots depict density of tagged SST-INs exhibiting cFos expression following fear conditioning **(E)** or morphine **(F)**. **(E)** Density of cFos+ tagged SST-INs: U = 0, p = 0.012, Mann-Whitney U test. **(F)** Density of cFos+ tagged SST-INs: U = 0, p = 0.030, Mann-Whitney U test. **(G-H)** Boxplot comparison of cFos+ cells within indicated brain regions after photoexcitation by t-test or Mann-Whitney U test, corrected for a false discovery rate of 0.1 by the Benjamini-Hochberg method. Basolateral amygdala, BLA; ventrolateral periaqueductal gray, vlPAG; paraventricular thalamus, PVT; lateral habenula, LHb; dorsomedial hypothalamus, DMH; dorsal midline thalamus, DMT; ventral

hippocampus area CA1, vCA1; dorsal hippocampus dentate gyrus, DG; caudate putamen, CPu; nucleus accumbens, NAc; dorsal PAG, dPAG; ventral tegmental area, VTA. **(G)** BLA: $t_8 = 3.34$, $p = 0.010$; vlPAG: $t_8 = 2.72$, $p = 0.026$; PVT: $t_8 = 3.34$, $p = 0.010$; LHb: $U = 0$, $p = 0.012$; DMH: $U = 0$, $p = 0.012$; DMT: $t_8 = 0.80$, $p = 0.45$; vCA1: $t_8 = 0.11$, $p = 0.92$; DG: $t_8 = 0.39$, $p = 0.71$; CPu: $t_8 = 0.69$, $p = 0.51$; NAc: $U = 10$, $p = 0.68$; dPAG: $t_8 = 0.83$, $p = 0.43$; VTA: $t_8 = 0.84$, $p = 0.42$. **(H)** BLA: $t_6 = 1.87$, $p = 0.11$; $U = 3$, $p = 0.19$; vlPAG: $t_6 = 2.72$, $p = 0.026$; PVT: $t_6 = 6.56$, $p = 5.99 \times 10^{-4}$; LHb: $t_6 = 1.15$, $p = 0.30$; DMH: $t_6 = 2.00$, $p = 0.092$; DMT: $t_6 = 0.73$, $p = 0.49$; vCA1: $U = 3$, $p = 0.19$; DG: $t_6 = 2.25$, $p = 0.065$; CPu: $U = 10$, $p = 0.66$; NAc: $t_6 = 7.42$, $p = 3.08 \times 10^{-4}$; dPAG: $t_6 = 9.46$, $p = 7.95 \times 10^{-5}$; VTA: $U = 0$, $p = 0.029$. * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$ by Tukey's post-hoc (**C**), Mann-Whitney U test (**E-H**), or unpaired two-sided t-test (**G-H**).

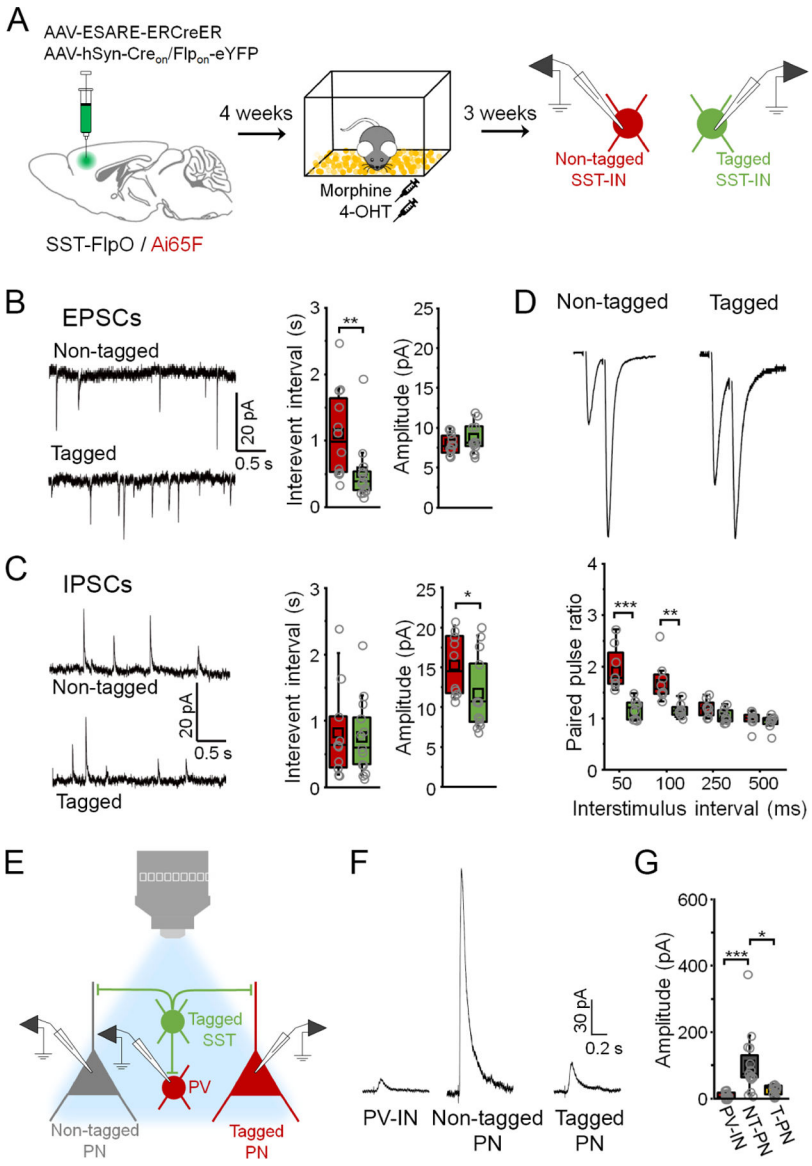


Figure 8. Synaptic connections formed by morphine-related SST-INs have distinct properties.

(A) Strategy for independent targeting of activated (eYFP+/ tdTomato+) versus nonactivated SST-INs (eYFP-/ tdTomato+) for electrophysiological recording following morphine treatment. **(B)** Boxplots depict spontaneous EPSC parameters from tagged (n = 13 cells) and non-tagged SST-INs (n = 12 cells) in the same slices (n = 4 slices from 4 mice). Interevent interval: U = 126, p = 0.0010, Mann-Whitney U-test. Amplitude: $t_{28} = -1.03$, p = 0.31, two-sided unpaired t-test. **(C)** Boxplots depict spontaneous IPSC parameters from tagged (n = 13 cells) and non-tagged SST-INs (n = 10 cells) in the same slices (n = 4 slices from 4 mice). Interevent interval: U = 89, p = 0.70, Mann-Whitney U-test. Amplitude: $t_{26} = 0.945$, p = 0.353, Mann-Whitney U-test. **(D)** Boxplots depict amplitude of EPSCs from tagged (n = 8 cells) and non-tagged SST-INs (n = 8 cells) in the same slices (n = 3 slices from 3 mice) during paired pulse stimulation. Paired pulse ratio: $F_{(3,21)} = 11.2$, $p = 1.36 \times 10^{-4}$, interaction between cell type and delay, 2-way repeated measures

ANOVA. **(E)** Strategy for electrophysiological analysis of connections from tagged SST-INs onto PV-INs (tdTomato+) as well as tagged (tdTomato+) and non-tagged PNs (tdTomato-) following morphine treatment. **(F)** Example IPSC traces. **(G)** Boxplots depict amplitude of IPSCs resulting from photoexcitation (460 nm, 1 ms, 0.1 Hz) of tagged SST-INs in PV-INs, non-tagged PNs (NT-PNs), and tagged PNs (T-PNs) following morphine treatment (n = 4 slices from 4 mice: PV-INs (n = 15 cells), NT-PNs (n = 16 cells), T-PNs (n = 12 cells)). $\chi^2 = 23.6$ (2), $p = 7.38 \times 10^{-6}$, Kruskal-Wallis ANOVA. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ by Mann-Whitney U test (**B, C**), Tukey's post-hoc test (**D**), or Dunn's post-hoc test (**G**). See also Figure S8.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-cFos	Millipore	Cat# abe457
Rabbit anti-somatostatin-14	Peninsula Laboratories	Cat# T-4103
Mouse anti-parvalbumin	Millipore	Cat# mab1572
Rabbit anti-vasoactive intestinal peptide	Immunostar	Cat# 20077
Goat anti-rabbit Alexa fluor 647	Jackson Immunoresearch	Cat# 111–605-045
Goat anti-mouse Alexa fluor 647	Jackson Immunoresearch	Cat# 115–605-003
Bacterial and Virus strains		
AAV1-EF1a-DIO-hChR2(H134R)-eYFP-WPRE	Karl Deisseroth Lab: Cre-activated AAV expression plasmids Unpublished	Addgene Cat# 20298
AAV1-CBA-FLEX-Arch-GFP	High-performance genetically targetable optical neural silencing by light-driven proton pumps. Chow BY, Han X, Dobry AS, Qian X, Chuong AS, Li M, Henninger MA, Belfort GM, Lin Y, Monahan PE, Boyden ES Nature. 2010 Jan 7. 463(7277):98–102.	Addgene Cat# 22222
AAV1-EF1a-DIO-eYFP-WPRE	Karl Deisseroth Lab: Double Floxed Inverted ORF Control Unpublished	Addgene Cat# 27056
AAV8-hSyn-mCherry	Karl Deisseroth Lab: mCherry Control Unpublished	Addgene Cat# 114472
rAAVDJ/nEF-Creon/Flpon-hChR2(H134R)-eYFP	University of North Carolina Gene Therapy Center Vector Core	n/a
rAAVDJ/nEF-Creon/Flpon-eYFP	University of North Carolina Gene Therapy Center Vector Core	n/a
AAV8-ESARE-ERCreER-PEST	Boston Children's Hospital Vector Core	n/a
Chemicals, peptides, and recombinant proteins		
4-hydroxytamoxifen (70% Z-isomer)	Sigma	Cat# H6278
Morphine sulfate	Sigma	Cat# 1448005
Tetrodotoxin citrate	Abcam	Cat# ab120055
4-Aminopyridine	Abcam	Cat# ab120122
Cyanquixaline disodium salt	Abcam	Cat# ab120044
Picrotoxin	Abcam	Cat# 120315
Experimental models: Organisms/strains		
Mouse: C57Bl/6J	Jackson Laboratory	Stock No. 000664
Mouse: SST-FlpO	Jackson Laboratory	Stock No. 028579
Mouse: PV-IRES-Cre	Jackson Laboratory	Stock No. 017320
Mouse: Ai9	Jackson Laboratory	Stock No. 007909
Mouse Ai65F	Jackson Laboratory	Stock No. 032864
Recombinant DNA		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pFBAAV-ESARE-ERCreER-PEST	Gift from Dr. Haruhiko Bito; Kawashima et al. (2013).	n/a
Software and algorithms		
MedAssociates Video Freeze	MedAssociates	https://www.med-associates.com/product/video-fear-conditioning/
Zen 2012 SP2	Zeiss	https://www.zeiss.com/microscopy/us/products/microscope-software/zen.html
Clampex 10.3.1	Molecular Devices	https://www.moleculardevices.com/products/axon-patch-clamp-system/acquisition-and-analysis-oftware/pclamp-software-suite
NIH ImageJ	National Institutes of Health	https://imagej.nih.gov/ij/index.html
VLC media player 2.1.3 Rincewind	VideoLAN	https://www.videolan.org/
MiniAnalysis	Synaptosoft	http://www.synaptosoft.com/MiniAnalysis/
Fusion v5.6 SuperFlex	Omnitech Electronics	http://www.omnitech-electronics.com/product/Fusion-Software/510
Photoshop 12.1	Adobe	https://www.adobe.com/products/photoshopfamily.html
OriginPro	OriginLab	https://www.originlab.com/
Prism 5	GraphPad	https://www.graphpad.com/scientific-software/prism
BioRender	BioRender	https://biorender.com/