HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2018 May 09.

Published in final edited form as:

Nat Neurosci. 2017 November ; 20(11): 1624–1633. doi:10.1038/nn.4651.

Cellular and Oscillatory Substrates of Fear Extinction Learning

Patrick Davis[1,2], **Yosif Zaki**[1], **Jamie Maguire**[1], and **Leon G. Reijmers**[1]

[1]Department of Neuroscience, Tufts University School of Medicine, Boston, MA, United States

[2]Medical Scientist Training Program and Graduate Program in Neuroscience, Sackler School of Graduate Biomedical Sciences, Tufts University, Boston, MA, United States

Abstract

The mammalian brain contains dedicated circuits for both the learned expression and suppression of fear. These circuits require precise coordination to facilitate the appropriate expression of fear behavior, but the mechanisms underlying this coordination remain unclear. Using a novel combination of chemogenetics, activity-based neuronal-ensemble labeling, and *in vivo* electrophysiology, we found that fear extinction learning confers parvalbumin-expressing (PV) interneurons in the basolateral amygdala (BLA) with a dedicated role in the selective suppression of a previously encoded fear memory and BLA fear-encoding neurons. In addition, following extinction learning, PV interneurons enable a competing interaction between a 6–12 Hz oscillation and a fear-associated 3–6 Hz oscillation within the BLA. Loss of this competition increases a 3–6 Hz oscillatory signature, with BLA→mPFC directionality signaling the recurrence of fear expression. The discovery of cellular and oscillatory substrates of fear extinction learning that critically depend on BLA PV-interneurons could inform therapies aimed at preventing the pathological recurrence of fear following extinction learning.

INTRODUCTION

The ability of neuronal circuits to integrate newly acquired information with memories of previous experiences in order to produce adaptive behavioral responses is critical to an animal's survival. This ability is exemplified by fear extinction learning, by which an animal reduces the expression of learned fear when repeatedly exposed to a stimulus or place that no longer predicts danger[1]. An analogous process in humans, known as exposure therapy, is a standard of care in the treatment of Post-Traumatic Stress Disorder and other anxiety disorders[2, 3]. Previous studies found that extinction learning does not lead to an erasure of the fear memory, but instead causes its suppression through the acquisition of a distinct memory[4]. The precise mechanisms by which the newly formed extinction memory can

Correspondence should be addressed to L.G.R. (leon.reijmers@tufts.edu).

AUTHOR CONTRIBUTIONS

P.D., J.M., and L.G.R conceived and designed the experiments. P.D. and Y.Z. executed the experiments. P.D., Y.Z., and L.G.R. analyzed the experiments. P.D. and L.G.R wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

interact with the original fear memory to suppress fear expression remain elusive[5]. We previously localized a neuronal correlate of a contextual fear memory in the basolateral amygdala (BLA) by using the TetTag transgenic mouse, which allows stable labeling of pyramidal neuron ensembles that are active during a defined temporal window[6]. In addition, we found that contextual fear extinction learning is associated with the silencing of these BLA fear neurons, as well as the target-specific remodeling of parvalbumin-positive perisomatic inhibitory synapses located around the BLA fear neurons[7]. These findings implicated BLA parvalbumin-positive interneurons (PV-interneurons) in the extinction-induced suppression of fear memories, which would reveal another function for BLA PV-interneurons in addition to their previously established role in fear learning[45]. We therefore decided to test if BLA PV-interneurons indeed suppress fear expression following extinction, and determine if BLA PV-interneurons mediate a direct interaction between an extinction memory circuit and a fear memory circuit.

RESULTS

BLA PV-interneurons selectively suppress conditioned freezing and fear ensemble activation following extinction

In order to investigate the causal role of BLA PV-interneurons in extinction-induced suppression of conditioned fear, we infused AAV-Syn-DIO-hM4Di-mCherry virus into the BLA of PV-Cre:Fos-tta:TetO-H2B-GFP triple-transgenic mice (PV-Cre:TetTag), thereby expressing the inhibitory DREADD (Design Receptor Exclusively Activated by Design Drug) receptor hM4Di selectively in BLA PV-interneurons in the background of TetTag mice (Fig. 1A, Fig. S1A–B). This permitted us to exert selective control over BLA PV-interneuron activity while tracking the effect on functionally relevant pyramidal neuron ensembles within the BLA[8]. Three weeks after virus infusion, mice were subjected to contextual fear conditioning and extinction, and neurons activated during fear conditioning or extinction were tagged with GFP (Fig. 1B, Fig. S1C). We previously found that tagged neurons in the BLA are excitatory pyramidal projection neurons, and that BLA interneurons are not tagged[7]. Mice exhibited robust freezing responses following conditioning and markedly suppressed freezing following extinction learning (Fig. S1D). In order to selectively test the role of BLA PV-interneurons in mediating the suppression of freezing following extinction, we measured freezing responses of mice in a post-extinction retrieval trial following injection of the DREADD ligand clozapine-N-oxide (CNO), thereby silencing BLA PV-interneurons. We found that CNO injection caused significantly more freezing in the conditioned context as compared to vehicle (VEH) injection (Fig. 1D, Fig. S1E–G). Interestingly, this effect was not observed in an unconditioned neutral context, indicating that CNO injection selectively impaired the ability to suppress conditioned freezing, rather than causing an overall increase in unconditioned fear (Fig. 1D). We reasoned that perhaps extinction learning confers the BLA PV-interneuron network with a specific role in fear suppression, which would be consistent with our previous structural plasticity findings[7]. Indeed, we found that silencing BLA PV-interneurons after fear conditioning, but without extinction learning, resulted in no behavioral effect, indicating that disinhibition *per se* does not lead to an increase in conditioned freezing behavior (Fig. 1E).

To identify ensemble-level mechanisms of this selective behavioral effect, we quantified the reactivation of neurons active during fear conditioning ("FC-tagged") within the BLA using TetTag (GFP) and the neuronal immediate-early gene Zif268 (ZIF) (Fig. 1B–C). Though this strategy is based on using two different immediate-early gene promoters for the detection of neuronal activation during encoding (transgenic Fos promoter) and retrieval (endogenous Zif promoter), we successfully used this promoter combination in previous studies to detect BLA ensembles associated with conditioned fear memories[6–7]. We found that silencing of PV-interneurons in the BLA resulted in increased reactivation of the FC-tagged GFP+ neurons (Fig. 1F, S1C$_1$). In contrast, PV-interneuron silencing had no effect on the activation of intermingled GFP− neurons (Fig. 1G), which should be noted are a relatively large and heterogeneous group of neurons that could include extinction neurons as well as fear neurons that were not GFP-tagged, or with GFP-tagging below detection threshold. The selective effect on GFP+ neurons indicated a selective disinhibition of the pyramidal neuron ensemble that was recruited during the encoding of the original fear memory. In support of this, the percentage of reactivated FC-tagged neurons correlated with the behavioral effect of CNO (Fig. 1F). One alternative explanation for these findings is that differences in intrinsic excitability bias a population of pyramidal neurons to be TetTagged with GFP, and would also make them more likely to be disinhibited by PV-interneuron silencing. To test this, we performed a similar experiment during which we TetTagged neurons active during extinction learning ("EXT-tagged", Fig. 1B), and quantified their reactivation during PV-interneuron silencing. We found that silencing BLA PV-interneurons had no effect on reactivation of EXT-tagged neurons, supporting the selective role of PV-interneurons in suppressing a BLA fear ensemble (Fig. 1H, Fig. S1C$_2$). Taken together, these results indicate that BLA PV-interneurons selectively suppress conditioned freezing behavior following extinction learning through the selective suppression of a neuronal correlate of the fear memory within the BLA.

A role for PV-interneurons in the selective suppression of a single BLA ensemble seems counterintuitive, given that each PV-interneuron in the BLA provides synaptic input to as many as 800–900 different BLA projection neurons[9]. Silencing BLA PV-interneurons would therefore be expected to disinhibit all projection neurons non-selectively. A potential substrate for this specificity is suggested by our previous finding that extinction learning causes a selective increase in perisomatic PV-positive synapses that are located around silent FC-tagged neurons[7]. If this mechanism indeed contributes to the selective suppression of the conditioned fear ensemble, then PV-interneuron silencing would have the largest disinhibiting effect on FC-tagged neurons receiving the largest amount of synaptic input from PV-interneurons. To test this prediction, we analyzed the relationship between perisomatic mCherry puncta, which correlated tightly with perisomatic PV puncta, and ZIF expression (Fig. 2A–B, Fig. S2A–B). We found that silencing of PV-interneurons following CNO injection led to an induction of a positive correlation between these two measures in FC-tagged neurons, demonstrating that FC-tagged neurons receiving the most perisomatic inhibition were the mostly likely to be disinhibited (Fig. 2C). Combined with the lack of correlation in the VEH group (Fig. 2C), this suggests that, following extinction, PV-interneurons exert a "normalizing" influence on FC-tagged neuron activation by more strongly inhibiting those FC-tagged neurons that would otherwise be the most active during

the retrieval trial. Interestingly, this normalizing function seems specific for FC-tagged neurons, as PV-interneuron silencing did not lead to positive correlations between perisomatic mCherry and ZIF in two other groups of BLA projection neurons: those not tagged during fear conditioning, and those tagged during extinction (Fig. 2D–E).

BLA PV-interneurons are necessary for a competing interaction between functionally opposed oscillations following extinction

Though we observed that silencing PV-interneurons preferentially disinhibited FC-tagged versus EXT-tagged neurons (Fig. 1F–H, Fig. 2C–E), we did not observe an overall difference in perisomatic mCherry between these two groups of BLA projection neurons (Fig. S2C). These data argue against the simple hypothesis that the selective role of PV-interneurons in suppressing the BLA fear ensemble is the result of BLA fear neurons as a whole receiving more perisomatic innervation from PV-interneurons. We therefore hypothesized that static synaptic differences alone might not fully account for the selective role of BLA PV-interneurons in post-extinction fear suppression, and that more dynamic circuit processes are also involved. To test this we decided to record local field potential (LFP) oscillations in the BLA that were previously shown to correlate with fear memory retrieval[10]. We recorded LFP oscillations in the BLA of freely behaving mice that expressed hM4Di in BLA PV-interneurons, and that were subjected to the same fear conditioning and extinction paradigm that was used for the TetTag experiments (Fig. 3A, Fig. S3A–B). We observed two distinct oscillations in the 3–12 Hz frequency band (Fig. 3B–C, Fig. S3C–D, Fig. S4). Interestingly, these two oscillations seemed to respond differently to PV-interneuron silencing, with an oscillation in the 3–6 Hz range increasing in power, while a different oscillation in the 6–12 Hz range decreased in power (figure 3C insert). We therefore decided to quantify the power of both oscillations across conditioning, extinction, and retrieval trials. We found that fear conditioning led to an increase in 3–6 Hz power (Fig. 3D–E), in agreement with a previous report[11]. Post-extinction silencing of PV-interneurons also increased the 3–6 Hz power, consistent with a role for this oscillation in producing freezing behavior (Fig. 3H–I). While fear conditioning had no effect on 6–12 Hz power, silencing PV-interneurons significantly reduced the power of this band (Fig. 3F,J). We therefore reasoned that perhaps the balance of these two oscillations determined the functional output of the BLA following extinction. Consistent with this notion, the power ratio of the two bands (3–6 Hz/6–12 Hz) correlated with the amount of freezing across and within mice following extinction (Fig. 3K, Fig. S3E–F). This correlation could merely reflect changes in oscillations whenever mice are involved in different behaviors (freezing versus non-freezing), without the oscillations actually contributing to behavior. Arguing against this, we found that restricting the analysis of the 3–6 Hz/6–12 Hz ratio during pure freezing bouts still revealed a CNO-induced increase in the 3–6 Hz/6–12 Hz ratio (figure S3G). Also, the 3–6 Hz/6–12 Hz ratio did not correlate with freezing before extinction (Fig. 3G, Fig. S3F). These findings suggest that extinction learning endows the balance of two different BLA oscillations with a novel function in the regulation of behavioral fear expression.

Since we observed that the power of the 3–6 and 6–12 Hz oscillations shifted in opposite directions following CNO injection, we decided to examine a potential dynamic interaction

between these two oscillations by cross-correlating their instantaneous power envelopes. We found that after extinction, but not before, the two sub-bands exhibited a significant negative correlation with one another, implying that extinction learning results in changes to the local circuitry that allow the two oscillations to compete with one another (Fig. 4A, C–D). Given our previous findings implicating structural plasticity of PV-interneurons in extinction learning[7], as well as the necessary role of PV-interneuron activity in fear suppression following, but not before, extinction learning (Fig. 1D–E), we reasoned that perhaps PV-interneurons played an integral role in the competing interaction between these two oscillations. In support of this, we found that silencing BLA PV-interneurons following extinction eliminated the negative correlation between the 3–6 Hz and 6–12 Hz oscillations, indicating that the mechanism segregating them is critically dependent on PV-interneuron activity (Fig. 4B–D). We did not detect a relationship between 6–12 Hz power and the 3–6 Hz : 6–12 Hz cross-correlation values (Fig. S5), making it unlikely that the elimination of the negative correlation following CNO injection was an artefact of reduced 6–12 Hz power (Fig. 3J). Taken together, these data imply that the post-extinction state is characterized by 6–12 Hz oscillations that interact with fear-associated 3–6 Hz oscillations through a mechanism dependent on BLA PV-interneurons.

BLA PV-interneurons gate information flow through reciprocal BLA-mPFC circuits following extinction

One possible explanation for the selective role of BLA PV-interneurons following extinction would be that the local PV-network is driven in part by upstream regions that relay learning-specific information critical to the suppression of fear. In order to identify inputs that might instruct the post-extinction role of the PV-network, we used Cre-dependent G-deleted rabies to specifically label monosynaptic inputs onto BLA PV-interneurons[12]. This revealed dense input from multiple brain regions, including two regions in the medial prefrontal cortex (mPFC): the infralimbic cortex (ILC) and the prelimbic cortex (PLC). The ILC, a region associated with extinction learning, appeared to contain a larger number of projection neurons that directly innervate BLA PV-interneurons as compared to the PLC, a region associated with fear learning and behavior (Fig. 5A)[1, 13]. To further investigate the mPFC-BLA circuitry, we used Cre-dependent rabies to selectively label the monosynaptic inputs onto neurons which project from BLA→mPFC. Interestingly, we observed mPFC projection neurons that synapse onto BLA→mPFC projecting neurons, thereby outlining a reciprocal loop between mPFC and BLA (Fig. 5B). Finally, we found that BLA→mPFC projecting neurons innervate both the PLC and ILC, with preferential targeting of superficial PLC and deep ILC (Fig. 5C). PV-interneurons are therefore synaptically positioned to gate information flow across a functionally relevant BLA-mPFC circuit, a feature potentially critical to their role in regulating fear behavior following extinction.

The observed circuit design implies that, in addition to exerting control over the activity of ensembles within BLA, PV-interneurons may also affect the activity of mPFC ensembles through dynamic modulation of mPFC-BLA reciprocal circuits. To determine whether BLA PV-interneurons contribute to post-extinction circuit dynamics within the mPFC, we examined the effect of silencing BLA PV-interneurons on LFP oscillatory activity in the mPFC using the same mice as used for the BLA analysis in Fig. 3 (Fig. 6A, Fig. S3A). We

found that silencing BLA PV-interneurons increased the 3–6 Hz/6–12 Hz power ratio in mPFC (Fig. 6B–D). The changes in the mPFC broadly mirrored the observed effects in BLA (Fig. 3I–K), indicating that BLA output to mPFC is regulated by PV-interneurons. If correct, then BLA PV-interneuron silencing should lead to changes in ensemble activation within the mPFC. To test this, we used the same TetTag mice as used for the BLA analysis in Fig. 1 (Fig. 6A). We found that silencing BLA PV-interneurons, thereby disinhibiting BLA fear neurons (Fig. 1F), increased the activation of TetTagged GFP+ neurons in superficial, but not deep, PLC (Fig. 6E, Fig. S1C$_3$). Interestingly, the reactivation of these PLC neurons correlated with freezing, indicating that the BLA and superficial PLC can cooperate to produce freezing behavior after extinction (Fig. 6F). In agreement with this, we found that superficial PLC neurons tended to be TetTagged in fear conditioned mice (Fig. 6G). In contrast to PLC, we found that silencing BLA PV-interneurons resulted in a selective decrease in the activity of TetTagged GFP+ neurons in the deep ILC (Fig. 6H, Fig. S1C$_4$). Further, we found that neurons in deep ILC were preferentially TetTagged in mice placed in a novel environment without footshock, as compared to mice left in home cage or mice that were fear conditioned, consistent with a putative role for deep ILC neurons in learned safety (Fig. 6I). These data support the notion that TetTagged neurons in the PLC and ILC play functionally distinct roles in the regulation of fear following extinction. Furthermore, these findings indicate that silencing PV interneurons in the BLA, thereby disinhibiting BLA fear neurons, might shift mPFC oscillatory activity through opposing changes in the activation of functionally distinct ensembles in the PLC and ILC.

Our tracing and functional data thus far point to a critical role for BLA PV-interneurons in gating information flow across the BLA-mPFC circuit in order to regulate the expression of fear following extinction. In further support of this, we found that silencing BLA PV-interneurons following extinction shifted BLA-mPFC coherence, a measure of interregional communication, towards the 3–6 Hz range (Fig. 7A). To examine whether changes in BLA oscillations can have a direct impact on mPFC oscillations, we assessed directionality across the BLA-mPFC circuit using a previously described method of cross-correlating instantaneous power envelopes (Fig. 7B)[14]. We found that freezing after fear conditioning, but not after extinction, tended to be accompanied by increased mPFC→BLA directionality in the 3–6 Hz band (Fig. S3H,I), in agreement with a previous report[11]. In contrast, we found that post-extinction freezing was characterized by a shift toward BLA→mPFC directionality in the 3–6 Hz band (Fig. 7C–D, Fig. S6), implying that although the circuit converges on a similar oscillatory state as assessed by a high 3–6 Hz/6–12 Hz power ratio, the pre- and post-extinction freezing states may arise from different circuit conditions. Further, we found that silencing BLA PV-interneurons increased the likelihood of BLA→mPFC directionality in the 3–6 Hz band, but reduced the likelihood of BLA→mPFC directionality in the 6–12 Hz band, further supporting that these two oscillations have distinct functions in the post-extinction state (Fig. 7E). These data indicate that BLA PV-interneurons route directional information transfer across BLA-mPFC circuits that are involved in the control of post-extinction fear expression.

DISCUSSION

Taken together, our data support a model whereby fear extinction learning induces remodeling of the BLA PV-interneuron network to allow competition between an extinction memory circuit and a fear memory circuit. This competition is represented in the divergent behavior of FC-tagged and EXT-tagged neurons within BLA, as well is in the negative correlation between the 3–6 Hz and 6–12 Hz oscillations. Following extinction, this competition causes suppression of BLA fear neurons and a fear-associated 3–6 Hz oscillation. Silencing BLA PV-interneurons after extinction disallows competition between the two circuits, thereby increasing the activation of the BLA fear neurons and the fear-associated 3–6 Hz oscillation, which is then signaled to the mPFC (Fig. 7F). This model has several important implications.

First, our findings reveal that a local PV-interneuron network can mediate a direct interaction between a new extinction memory and a previous fear memory, which are represented by distinct network states. Recent studies have demonstrated that PV-networks can adopt multiple learning-dependent network states that can modulate synaptic plasticity and circuit output[7, 15–17]. Here we report that extinction learning confers PV-interneurons in the BLA with a selective role in the suppression of conditioned fear behavior, and in the silencing of a neuronal fear memory correlate within the BLA. Our findings are consistent with our previous discovery of an extinction-induced target-specific structural plasticity of BLA PV-interneurons that likely results in increased perisomatic inhibition of FC-tagged neurons[7]. Target-specific changes in synapse numbers and/or strength are likely to alter the balance of local “winner-take-all” competitive inhibitory networks, thereby biasing the control of circuit output in favor of a select population of pyramidal neurons at the expense of others[18, 19]. In agreement with a role for PV-interneurons in mediating this competition, silencing PV-interneurons during extinction retrieval alters the activity of BLA fear neurons, but not of other functional ensembles within BLA. The ability of BLA PV-interneurons to dynamically regulate the activation of BLA projection neurons is likely supported by their recruitment through both feedback input from BLA projection neurons, as well as feedforward input from projection neurons outside of the BLA, including mPFC and ventral hippocampus[20, 35]. Divergent feedback input from BLA PV-interneurons might contribute to lateral inhibition between anatomically and functionally distinct populations of BLA projection neurons, as was recently observed[21]. In addition, our tracing data indicate the participation of BLA PV-interneurons in reciprocal connections between mPFC and BLA[22, 23]. Therefore, BLA PV-interneurons could indirectly modulate their own activity via engagement of reciprocal connections with mPFC and feedforward inhibition. It remains to be determined what effect the BLA-mPFC reciprocal connections have on BLA circuit dynamics, and whether input from mPFC is necessary to confer the BLA PV-network with its post-extinction fear-suppressing function. Also, it will be of interest to investigate how the role of the BLA PV-network in post-extinction fear suppression might complement other inhibitory circuits that have been implicated in post-extinction fear suppression, such as the intercalated neurons[36–40], and if the involvement of the BLA PV-network in post-extinction fear suppression is similar for contextual and cued fear[5].

Second, we propose that while the 3–6 Hz oscillation in the BLA-mPFC circuit consistently correlates with freezing behavior[11], the valence of the 6–12 Hz oscillation, and indeed its relationship with the 3–6 Hz oscillation, is dependent on previous learning. The 6–12 Hz oscillation resembles classic hippocampal theta, and hippocampal activity can contribute to both fear expression and fear suppression, indicating a lack of predefined valence[24]. Accordingly, hippocampal neurons active during fear learning could be manipulated to change their valence through additional learning, while similar neurons in BLA could not[25]. A recent study demonstrated that direct ventral hippocampus (VH) input to BLA is required for the expression of contextual fear, but a role for this pathway in extinction learning remains to be determined[26]. Since VH→BLA projections directly innervate BLA PV-interneurons, and can thereby modulate synaptic plasticity[27, 28, 35], it is possible that in the behaving animal VH→BLA projections can have opposing effects on amygdalar activity and behavioral output depending on previous learning and the state of the BLA PV-interneuron network. It should be noted that the VH can also regulate the BLA microcircuit indirectly through projections to mPFC neurons that project to the BLA[41–42]. Further studies will be necessary to determine the origin and precise function of the BLA 6–12 Hz oscillation, as well as how this function is dependent on the BLA PV-network.

Third, the increased reactivation of fear neurons within BLA coincided with a shift toward 3–6 Hz oscillations, suggesting a causal relationship between the 3–6 Hz oscillation and BLA fear neuron activation. One possible explanation would be that BLA fear neurons act as resonators that are preferentially tuned to the 3–6 Hz frequency band, or conversely they could actively participate in the production of the 3–6 Hz oscillation through engagement of local circuit mechanisms[29]. While BLA does not appear to generate low-frequency oscillations in slice preparations, previous studies have found that projection neurons within the BLA have resonance properties aligned to the 3–6 Hz range[30]. An interesting avenue for future studies would be to explore whether a learning experience can modulate intrinsic resonance properties of select neuronal ensembles, thereby entraining them to a particular functionally relevant frequency band. If correct, then the artificial induction of a 4 Hz or 8 Hz oscillation in BLA following fear conditioning and extinction might lead to the selective recruitment of fear or extinction neurons, respectively[21, 31]. Such an outcome would be consistent with the notion that the 3–12 Hz LFP oscillations recorded in the BLA reflect oscillatory activity of BLA neurons[46–50]. However, it is important to point out that pending future confirmation that BLA neurons are phase-locked to 3–12 Hz oscillations, or gain-of-function manipulations of 3–12 Hz oscillatory activity within the BLA, we can not rule out that the observed 3–12 Hz oscillations in the BLA resulted from volume conduction of oscillatory activity that occurred outside of the BLA.

Fourth, the observed relationship between LFP oscillations and freezing suggest that these oscillations make a causal contribution to freezing. Indeed, artificial induction of a 4 Hz oscillation in dorsal mPFC was found sufficient to synchronize mPFC single units and induce freezing behavior in mice[11]. It should be noted that although our correlative data suggest that, similar to the mPFC 4 Hz oscillation, a 3–6 Hz oscillation in the BLA could cause freezing behavior, this remains a hypothesis that will have to be confirmed in future studies that employ frequency-specific manipulations of BLA activity. Until then, we cannot

exclude the possibility that at least some of the observed oscillatory changes in our study resulted from changes in behavior, instead of the other way around.

Finally, our findings indicate that a state of fear that re-emergences after extinction can be distinct from the state of fear before extinction. In contrast to the pre-extinction state, in which mPFC activity can drive BLA during fear behavior[11], our findings indicate that the return of extinguished fear can be caused by a signal from the BLA to the mPFC. Based on our data, we propose that this BLA→mPFC signal shifts mPFC activity toward a pro-fear state by causing opposite changes in the activation state of PLC and ILC ensembles that are associated with promoting and suppressing fear, respectively[1, 13]. Accordingly, recent studies reported a role for BLA→mPFC projections in the retention and extinction of fear memories[44], and reported the occurrence of a BLA→PLC signal during fear states that are caused by conflicting cues[43]. The conditioned stimulus in our behavioral protocol (i.e. context) could be considered a conflicting cue, as it was previously associated with both danger (fear conditioning) and safety (extinction). Our observation of a BLA→mPFC signal during post-extinction fear is also consistent with previous reports of latent fear circuits in the human amygdala that can contribute to the return of fear[2, 31–33]. A better mechanistic understanding of the conditions that lead to the return of fear behavior following extinction, such as during spontaneous recovery, would greatly aid the development of better treatments for anxiety disorders such as Post-Traumatic Stress Disorder. Our findings suggest that a therapy or manipulation that reduces the likelihood of BLA fear neuron reactivation, 3–6 Hz oscillatory activity, and/or BLA→mPFC engagement could be beneficial for preventing the return of fear following extinction learning.

ONLINE METHODS

Animals

All animal procedures were performed in accordance with the NIH Health Guide for the Care and Use of Laboratory Animals and were approved by the Tufts University Institutional Animal Care and Use Committee. The TetTag mice used in this study were heterozygous for two transgenes: c-fos promoter-driven tetracycline transactivator (cfosP-tTa) and a tet operator-driven fusion of histone2B and eGFP (tetO-His2BGFP). PV-Cre and PV-Cre:TetTag mice used in this study were heterozygous for a PV-IRES-Cre knock-in locus (B6; 129p2-Pvalb$^{tm1(cre)Arbr}$/J). Mice had food and water ad libitum and were socially housed until the start of behavioral experiments, which was at an age of at least 12 weeks. Mice were kept on regular light-dark cycle, and all experiments were performed during the light phase.

Stereotaxic surgery

Mice were anesthetized with isoflurane, held in a stereotaxic apparatus (Kopf), and injected with virus. After injection, the needle was left in place for 10 min before slowly retracting. The incision was sutured, and mice were weighed and monitored to ensure recovery. For DREADD experiments, 400nl of AAV-Syn-DIO-hm4Di-mCherry (UNC Vector Core, Bryan Roth) was injected bilaterally into basolateral amygdala (BLA; AP −1.35 mm, ML ± 3.45 mm, DV −5.15 mm). For anterograde tracing of BLA→mPFC projections, 200nl of AAV9-

CaMKII-Cre-GFP was injected into mPFC (AP +0.75, ML ± 0.3mm, DV −2.1 mm) and 400nl of AAV-Syn-DIO-hm4Di-mCherry was injected into the BLA. For rabies tracing in PV-Cre mice, 250nl of AAV-EF1a-FLEX-GTB "Helper virus" was injected into the BLA, and glycoprotein deleted rabies virus (EnvA)SAD-ΔG-mCherry (generously provided by Dr. Ed Callaway and the Salk Institute Vector Core), was mixed with CTB647 for injection site localization and injected 3 weeks later into the same region. For retrograde tracing of all input to BLA, 75nl of CTB555 was injected into the BLA. Mice were sacrificed and tissue was collected and processed 5–9 days following rabies injection. All coordinates are relative to bregma.

Behavior

3–6 month old PV-Cre and PV-Cre:TetTag mice were used for the study. None of the mice had prior procedures or testing performed. Mice were randomly assigned to groups, except for ensuring equal distribution of females/males across groups. Method of randomization was based on arbitrary numbers that were assigned at the time of weaning. The design of experiment 1 is summarized in Fig. 1B. Mice were subjected to contextual fear conditioning consisting of three training trials (FC1, FC2, and FC3) with 3 hours between each trial. The total duration of each training trial was 500 s. A training trial started with placing the mouse in a square chamber with grid floor (context A) (Coulbourn Instruments; H10-11RTC, 120W 3 100D 3 120H). At 198 s, 278 s, 358 s, and 438 s, a foot shock was delivered (2 s, 0.70 mA). On days 2 and 3 (or 4 and 5 for the EXT-tagged group), mice were subjected to four extinction trials per day. Each extinction trial lasted 1200s with a trial interval of 2 hr. For each extinction trial, mice were placed in the same box used for fear conditioning without receiving foot shocks. On days 4 and 5, mice were tested over 500 s during a single retrieval test in context A. A subset of mice in Experiment 1 also were tested in a neutral context without footshock, which consisted of a square plastic box with bedding sprayed with 10% acetic acid and striped walls. CNO/VEH mice received an IP injection of 8–10mg/kg clozapine-N-oxide (CNO) on day 4 and of vehicle (VEH; 5% DMSO in saline) on day 5. VEH/CNO received injections in the reverse order. CNO/VEH and VEH/CNO mice were counterbalanced within each experiment. All TetTag mice were perfused 90 min after the final retrieval trial (day 5 for FC-tagged, day 7 for EXT-tagged).

For labeling of activated cells, TetTag mice were raised on food with doxycycline (40 mg doxycycline/kg chow). For the FC-tagged group, 4 days before conditioning, mice were individually housed, and doxycycline was removed from the food. After the last fear conditioning trial on day 1, mice were put on food with a high dose of doxycycline (1 g/kg) to rapidly block the tagging of neurons activated after fear conditioning. On day 2 mice were put back on the regular dose of doxycycline (40 mg/kg). For the EXT-tagged group, mice were conditioned while on doxycycline in an identical fashion to the FC-tagged group. Doxycycline was subsequently removed from food and animals remained in homecage for 2 days before beginning two days of extinction trials. Following the last extinction trial on day 5, mice were put on food with high doxycycline (1g/kg).

For homecage versus no-shock versus fear conditioned comparisons, a separate cohort of TetTag mice was split into 3 groups. Homecage mice were taken off dox for four days and

left in their homecage. Fear conditioned mice were taken off dox for four days prior to contextual fear conditioning as described above, then placed back on high doxycycline chow as described above. No-shock mice were placed in the conditioning context for three identical trials with the exception that no shock was given. No-shock and fear conditioned mice were perfused 90 min following a retrieval trial that took place 3 days after no-shock or fear conditioning. Homecage mice were perfused at the same time.

Quantification of freezing

Freezing behavior was measured using a digital camera connected to a computer with Actimetrics FreezeFrame software. The minimum bout length was 1.5 s and the threshold for freezing behavior was set prior to the onset of the experiment, and was the same for all subsequent trials for each animal. Freezing scores were obtained by averaging freezing during the entire trial, unless otherwise indicated.

Tissue preparation and immunohistochemistry

Ninety minutes after retrieval, mice were deeply anesthetized with ketamine/xylazine and intracardially perfused with 0.1 M phosphate buffer (PB) followed by 4% paraformaldehyde (PFA 4%) dissolved in 0.1 M PB. Brains were extracted and post-fixed in PFA 4% for 24 hr. Brains were transferred to 30% sucrose for 48–72 hours before slicing 40μm (for hM4Di-mCherry localization and Nissl staining) or 25 μm (for all immunofluorescence) coronal sections of the entire brain using a cryostat. Sections were stored in phosphate-buffered saline (PBS) with 0.025% sodium-azide at 4°C until use. For immunofluoroscent staining, sections were blocked for 1 hr at room temperature in PBS-T (PBS with 0.25 % Triton X-100) with 8% normal goat serum. Sections were incubated in rabbit anti-Zif268 (Santa-Cruz sc-189; polyclonal; 1:3,000) with mouse anti-PV (Millipore MAB1572; monoclonal; 1:2,000), or rabbit anti-RFP (Rockland 600-401-379; polyclonal; 1:1500) at 4°C for 48–72 hours. Secondary antibodies (Jackson ImmunoResearch; goat anti-rabbit 549 1/1,500, goat anti-mouse 647 1/500) were diluted in the blocking solution and were then applied to the sections for 2 hours at room temperature followed by three rinses for 15 min in PBS-T. Sections were mounted on slides and coverslipped after a brief wash with 0.00005% DAPI in PBS-T to label cell nuclei and stored at 4°C.

Microscopy

A wide-field epifluorescent microscope (Keyence BZ-X700) was used to acquire images for electrode and injection site validation, and for TetTag data excluding perisomatic mCherry analysis. 10–20× images were obtained and stitched together using Keyence software. Acquisition settings were optimized for each brain region and were identical across groups. A minimum of 8 total amygdalae or 6 other brain region sections (4 or 3 bilateral) per animal was analyzed, after excluding sections for quality reasons. Sections including damage from injection site were excluded. A confocal laser-scanning microscope (Nikon A1R, or Leica SPE) was used to acquire images for mCherry/PV overlap analysis and mCherry-Zif correlation analysis. The settings for PMT, laser power, gain, and offset were identical between experimental groups. For mCherry/PV overlap, 20× z-stacks were acquired and the maximum intensity projection was used for analysis. For mCherry-Zif

correlation analysis, 20× images were used. A minimum of 4 sections containing the amygdala per animal was used for this analysis.

Quantification of activated cells

ImageJ software was used to select and count the total number of DAPI-, GFP-, and Zif-positive nuclei and nuclei double positive for GFP and Zif. In order to avoid bias, all three cell types (GFP+Zif−, GFP+Zif+, GFP−Zif+) were selected from the same pictures, and the threshold settings for GFP and Zif were identical across all mice.

Perisomatic analysis

Selection of GFP-labeled cells for confocal perisomatic analysis was designed to only include excitatory neurons as described previously[7]. mCherry signal was distinguished from nuclear ZIF using ImageJ image calculator function and DAPI as a mask for nuclear signal. Absolute values of fluorescence were calculated for each individual neuron in an unbiased way without knowledge of GFP+ fluorescence or experimental group. Values were then normalized within each section to reduce noise from imaging and tissue processing and to allow for comparison across animals.

Electrophysiology

All LFP data were acquired using pre-fabricated headmounts (Pinnacle 8201) for 2EEG/1EMG recordings. The headmounts were affixed to the skull with stainless steel screws which also act as EEG reference and ground electrodes. Stainless steel wires served as the EEG electrodes, which were placed into the BLA and mPFC (BLA AP −1.35 mm, ML ± 3.45 mm, DV −5.15 mm, mPFC AP +0.75, ML ± 0.3mm, DV −2.1 mm). Local field potentials were recorded using a 100× preamplifier (Pinnacle 8202-SE) at a sampling rate of 4 kHz. The data were analyzed and filtered offline using LabChart software.

All power spectra quantification data were generated using LabChart's DataPad function, using the first 4 minutes of each behavioral trial. Normalized power and 3–6/6–12 Hz ratio (Fig. 3, S3) were calculated by subtracting the power or power ratio value for the first 2 minutes of the first conditioning trial, during which the animal is acclimating to the context, from the ratio for the CNO or VEH retrieval trials.

Directionality was calculated using a modified process as described by others. In short, instantaneous amplitudes of filtered traces from mPFC and BLA were generated in LabChart. Files were then exported into ClampFit and cross-correlated after being segmented into 20s bins, and statistics were calculated using the population distribution of lags per individual trial (based on the first 4 minutes of each trial). For finer temporal scale directionality analysis, periods of at least 10 seconds of motion followed by at least 10 seconds of freezing (one per animal per trial) were selected to maximize signal to noise. The onset of freezing was demarked 0s and directionality analysis was performed from −5 to 5 sec in 1 sec bins with 0.5 sec steps. Lags were then averaged across animals on a per-second basis for visualization, or, for quantification, for the 5s prior to freezing and the 5s after the onset of freezing. For pure freezing analysis, separate files were created in LabChart collating all periods of pure freezing or non-freezing for a given trial using second-by-

second freezing data from FreezeView. These files were then subsequently used for analysis as indicated.

Sub-band cross-correlation analysis was performed in ClampFit by cross-correlating instantaneous amplitudes of 3–6 Hz and 6–12 Hz bands over a 2 min window at beginning of trial and per each trial per animal. Statistics were performed using area under curve of cross-correlogram. Onset of freezing analysis was performed using 1 sec bins from the same freezing onset epochs as described above. Spectrograms and coherence data were generated using the EEGLAB plugin in MATLAB (spectrograms: 3 cycle wavelets with a Hanning-taper window; coherence: pop_crossf function)[51]. Quantification of peak coherence was performed using the first 4 minutes of each trial, and exporting coherence curves to GraphPad Prism for analysis.

For Granger causality analysis, linear detrending and normalization were performed on the LFP signals prior to analysis, and the SIFT toolbox in EEGLAB was used to fit a higher order vector autoregressive model to the processes. Data were tested for stability in time and model order was determined using the Akaike information criterion. The SIFT toolbox in EEGLAB was used to calculate normalized Directed Transfer Function (nDTF) values for BLA→mPFC and mPFC→BLA using the first 2 minutes of retrieval trials (for CNO versus vehicle comparison), or using 10 second bins of periods of high or low freezing (for freezing versus no freezing comparison).

Exclusion of mice from analysis

Mice were excluded from analysis if they did not show effective fear conditioning (both FC3 and EXT1 freezing less than 35%, or EXT1 freezing less than 60% of FC3 freezing: n=3 TetTag mice excluded, n=0 LFP mice excluded), or effective fear extinction (EXT8 freezing more than 50% of EXT1 freezing: n=3 TetTag mice excluded, n=0 LFP mice excluded), or if AAV-Syn-DIO-hm4Di-mCherry infusion was not targeted to the BLA (n=4 TetTag mice excluded, n=4 LFP mice excluded). TetTag mice were excluded from ensemble activation analysis if the average number of GFP+ cells in the each BLA section was less than 10 GFP + cells (n=2 TetTag mice excluded). Mice were excluded from LFP analysis if electrode placement was outside of the BLA (n=3 LFP mice excluded). This left a total of 11 mice for LFP analysis, of which 10 mice also had correct mPFC placement. Two out of the 11 mice with correct BLA electrode placement were excluded from the POST-FC cross-correlogram analysis due to insufficient quality of the LFP data (Fig. 4). Of the 10 mice with correct BLA and mPFC electrode placement, one mouse had low-frequency noise in the recordings that interfered with the power analysis, but not with the directionality analysis. This mouse was therefore only used for directionality analysis, resulting in an n=9 for power analysis (Fig. 6), and an n=10 for directionality analysis (Fig. 7). For directionality analysis around freezing onset, a period of at least 10 seconds of motion followed by at least 10 seconds of freezing was used to maximize signal to noise, and mice without such an epoch during the retrieval trial were therefore excluded from the analysis. This resulted in a higher exclusion rate for the VEH trials, since VEH-injected mice had lower overall freezing levels (Fig. 7: trials from n=5–6 VEH-injected mice excluded, trials from n=2–3 CNO-injected mice excluded).

Statistics

No statistical methods were used to pre-determine sample sizes, but sample sizes are similar to those reported in previous publications[6, 10]. Data collection and analysis were not performed blind to the conditions of the experiments. Statistical tests were performed using Prism (GraphPad) and are indicated in the figure legends. All statistical tests were two-tailed. Assumptions about normal distributions were tested using the Shapiro-Wilk normality test, and assumptions about equality of variances were tested using an F-test. Non-parametric tests were used when assumptions for the parametric test were not met (unpaired t-test: normal distribution and equal variance of population values; paired t-test: normal distribution of differences; linear regression: normal distribution of residuals). All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

Life sciences reporting summary

A summary of experimental design and software used is also provided in the Life Sciences Reporting Summary.

Data availability

All relevant data supporting the findings of this study are available from the corresponding author upon reasonable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank B. Roth (UNC vector core) and E. Callaway (Salk Institute) for reagents. We thank A. Poulopoulos and T. Papouin for discussions and critical reading of the manuscript. We thank J. Sasaki Russell and S. Viola for technical assistance. This work was supported in part by grants to L.G.R. (NIH R01 MH104589) and J.M. (NIH R01 NS102937), and by the Tufts Center for Neuroscience Research (NIH P30 NS047243). P.D. was supported by the Synapse Neurobiology Training Program (NIH T32 NS061764), and the Medical Scientist Training Program at Tufts University (NIH T32 GM008448).

References

1. Milad MR, Quirk GJ. Fear extinction as a model for translational neuroscience: ten years of progress. Annu Rev Psychol. 2012; 63:129–151. [PubMed: 22129456]
2. Kearns MC, Ressler KJ, Zatzick D, Rothbaum BO. Early interventions for PTSD: a review. Depress Anxiety. 2012; 29:833–842. [PubMed: 22941845]
3. Foa EB. Prolonged exposure therapy: past, present, and future. Depress Anxiety. 2011; 28:1043–1047. [PubMed: 22134957]
4. Myers KM, Davis M. Mechanisms of fear extinction. Mol Psychiatry. 2007; 12:120–150. [PubMed: 17160066]
5. Tovote P, Fadok JP, Luthi A. Neuronal circuits for fear and anxiety. Nat Rev Neurosci. 2015; 16:317–331. [PubMed: 25991441]
6. Reijmers LG, Perkins BL, Matsuo N, Mayford M. Localization of a stable neural correlate of associative memory. Science. 2007; 317:1230–1233. [PubMed: 17761885]

7. Trouche S, Sasaki Jennifer M, Tu T, Reijmers Leon G. Fear Extinction Causes Target-Specific Remodeling of Perisomatic Inhibitory Synapses. Neuron. 2013; 80:1054–1065. [PubMed: 24183705]
8. Armbruster BN, Li X, Pausch MH, Herlitze S, Roth BL. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. Proceedings of the National Academy of Sciences. 2007; 104:5163–5168.
9. Vereczki VK, et al. Synaptic Organization of Perisomatic GABAergic Inputs onto the Principal Cells of the Mouse Basolateral Amygdala. Frontiers in Neuroanatomy. 2016; 10:20. [PubMed: 27013983]
10. Seidenbecher T, Laxmi TR, Stork O, Pape HC. Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science. 2003; 301:846–850. [PubMed: 12907806]
11. Karalis N, et al. 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. Nat Neurosci. 2016; 19:605–612. [PubMed: 26878674]
12. Wall NR, Wickersham IR, Cetin A, De La Parra M, Callaway EM. Monosynaptic circuit tracing in vivo through Cre-dependent targeting and complementation of modified rabies virus. Proceedings of the National Academy of Sciences. 2010; 107:21848–21853.
13. Giustino TF, Maren S. The Role of the Medial Prefrontal Cortex in the Conditioning and Extinction of Fear. Front Behav Neurosci. 2015; 9:298. [PubMed: 26617500]
14. Adhikari A, Sigurdsson T, Topiwala MA, Gordon JA. Cross-correlation of instantaneous amplitudes of field potential oscillations: A straightforward method to estimate the directionality and lag between brain areas. Journal of Neuroscience Methods. 2010; 191:191–200. [PubMed: 20600317]
15. Donato F, Rompani SB, Caroni P. Parvalbumin-expressing basket-cell network plasticity induced by experience regulates adult learning. Nature. 2013; 504:272–276. [PubMed: 24336286]
16. Lucas, Elizabeth K., Jegarl, Anita M., Morishita, H., Clem, Roger L. Multimodal and Site-Specific Plasticity of Amygdala Parvalbumin Interneurons after Fear Learning. Neuron. 2016; 91:629–643. [PubMed: 27427462]
17. Rashid AJ, et al. Competition between engrams influences fear memory formation and recall. Science. 2016; 353:383–387. [PubMed: 27463673]
18. de Almeida L, Idiart M, Lisman JE. A second function of gamma frequency oscillations: an E%-max winner-take-all mechanism selects which cells fire. J Neurosci. 2009; 29:7497–7503. [PubMed: 19515917]
19. Roux L, Buzsáki G. Tasks for inhibitory interneurons in intact brain circuits. Neuropharmacology. 2015; 88:10–23. [PubMed: 25239808]
20. Smith Y, Paré J-F, Paré D. Differential innervation of parvalbumin-immunoreactive interneurons of the basolateral amygdaloid complex by cortical and intrinsic inputs. The Journal of Comparative Neurology. 2000; 416:496–508. [PubMed: 10660880]
21. Kim J, Pignatelli M, Xu S, Itohara S, Tonegawa S. Antagonistic negative and positive neurons of the basolateral amygdala. Nat Neurosci. 2016 advance online publication.
22. Little JP, Carter AG. Synaptic mechanisms underlying strong reciprocal connectivity between the medial prefrontal cortex and basolateral amygdala. J Neurosci. 2013; 33:15333–15342. [PubMed: 24068800]
23. McGarry LM, Carter AG. Inhibitory Gating of Basolateral Amygdala Inputs to the Prefrontal Cortex. The Journal of Neuroscience. 2016; 36:9391–9406. [PubMed: 27605614]
24. Maren S, Phan KL, Liberzon I. The contextual brain: implications for fear conditioning, extinction and psychopathology. Nat Rev Neurosci. 2013; 14:417–428. [PubMed: 23635870]
25. Redondo RL, et al. Bidirectional switch of the valence associated with a hippocampal contextual memory engram. Nature. 2014; 513:426–430. [PubMed: 25162525]
26. Xu C, et al. Distinct Hippocampal Pathways Mediate Dissociable Roles of Context in Memory Retrieval. Cell. 2016; 167:961–972. [PubMed: 27773481]
27. Bazelot M, et al. Hippocampal Theta Input to the Amygdala Shapes Feedforward Inhibition to Gate Heterosynaptic Plasticity. Neuron. 2015; 87:1290–1303. [PubMed: 26402610]

28. Bienvenu, Thomas CM., Busti, D., Magill, Peter J., Ferraguti, F., Capogna, M. Cell-Type-Specific Recruitment of Amygdala Interneurons to Hippocampal Theta Rhythm and Noxious Stimuli In Vivo. Neuron. 2012; 74:1059–1074. [PubMed: 22726836]
29. Buzsáki G. Theta Oscillations in the Hippocampus. Neuron. 2002; 33:325–340. [PubMed: 11832222]
30. Ryan SJ, et al. Spike-Timing Precision and Neuronal Synchrony Are Enhanced by an Interaction between Synaptic Inhibition and Membrane Oscillations in the Amygdala. PLoS One. 2012; 7:e35320. [PubMed: 22563382]
31. Herry C, et al. Switching on and off fear by distinct neuronal circuits. Nature. 2008; 454:600–606. [PubMed: 18615015]
32. Schiller D, et al. Preventing the return of fear in humans using reconsolidation update mechanisms. Nature. 2010; 463:49–53. [PubMed: 20010606]
33. Agren T, et al. Disruption of Reconsolidation Erases a Fear Memory Trace in the Human Amygdala. Science. 2012; 337:1550–1552. [PubMed: 22997340]
34. Björkstrand J, et al. Disrupting Reconsolidation Attenuates Long-Term Fear Memory in the Human Amygdala and Facilitates Approach Behavior. Current Biology. 2016; 26:2690–2695. [PubMed: 27568591]
35. Hübner C, Bosch D, Gall A, Lüthi A, Ehrlich I. Ex vivo dissection of optogenetically activated mPFC and hippocampal inputs to neurons in the basolateral amygdala: implications for fear and emotional memory. Frontiers in Behavioral Neuroscience. 2014; 8:64. [PubMed: 24634648]
36. Quirk GJ, Likhtik E, Pelletier JG, Pare D. Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. J Neurosci. 2003; 23:8800–8807. [PubMed: 14507980]
37. Royer S, Pare D. Bidirectional synaptic plasticity in intercalated amygdala neurons and the extinction of conditioned fear responses. Neuroscience. 2002; 115:455–462. [PubMed: 12421611]
38. Likhtik E, Popa D, Apergis-Schoute J, Fidacaro GA, Pare D. Amygdala intercalated neurons are required for expression of fear extinction. Nature. 2008; 454:642–645. [PubMed: 18615014]
39. Amano T, Unal CT, Pare D. Synaptic correlates of fear extinction in the amygdala. Nat Neurosci. 2010; 13:489–494. [PubMed: 20208529]
40. Jungling K, et al. Neuropeptide S-mediated control of fear expression and extinction: role of intercalated GABAergic neurons in the amygdala. Neuron. 2008; 59:298–310. [PubMed: 18667157]
41. Orsini CA, Kim JH, Knapska E, Maren S. Hippocampal and prefrontal projections to the Basal amygdala mediate contextual regulation of fear after extinction. J Neurosci. 2011; 31:17269–17277. [PubMed: 22114293]
42. Sotres-Bayon F, Sierra-Mercado D, Pardilla-Delgado E, Quirk Gregory J. Gating of Fear in Prelimbic Cortex by Hippocampal and Amygdala Inputs. Neuron. 2012; 76:804–812. [PubMed: 23177964]
43. Burgos-Robles A, et al. Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat Neurosci. 2017; 20:824–835. [PubMed: 28436980]
44. Klavir O, Prigge M, Sarel A, Paz R, Yizhar O. Manipulating fear associations via optogenetic modulation of amygdala inputs to prefrontal cortex. Nat Neurosci. 2017; 20:836–844. [PubMed: 28288126]
45. Wolff SBE, et al. Amygdala interneuron subtypes control fear learning through disinhibition. Nature. 2014; 509:453–458. [PubMed: 24814341]
46. Pare D, Gaudreau H. Projection cells and interneurons of the lateral and basolateral amygdala: distinct firing patterns and differential relation to theta and delta rhythms in conscious cats. J Neurosci. 1996; 16:3334–3350. [PubMed: 8627370]
47. Popa D, Duvarci S, Popescu AT, Léna C, Paré D. Coherent amygdalocortical theta promotes fear memory consolidation during paradoxical sleep. Proceedings of the National Academy of Sciences. 2010; 107:6516–6519.
48. Bienvenu, Thomas CM., Busti, D., Magill, Peter J., Ferraguti, F., Capogna, M. Cell-Type-Specific Recruitment of Amygdala Interneurons to Hippocampal Theta Rhythm and Noxious Stimuli In Vivo. Neuron. 2012; 74:1059–1074. [PubMed: 22726836]

49. Likhtik E, Stujenske JM, Topiwala MA, Harris AZ, Gordon JA. Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nat Neurosci. 2014; 17:106–113. [PubMed: 24241397]

50. Stujenske JM, Likhtik E, Topiwala MA, Gordon JA. Fear and safety engage competing patterns of theta-gamma coupling in the basolateral amygdala. Neuron. 2014; 83:919–933. [PubMed: 25144877]

51. Delorme A, et al. EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. Comput Intell Neurosci. 2011; 2011:130714. [PubMed: 21687590]

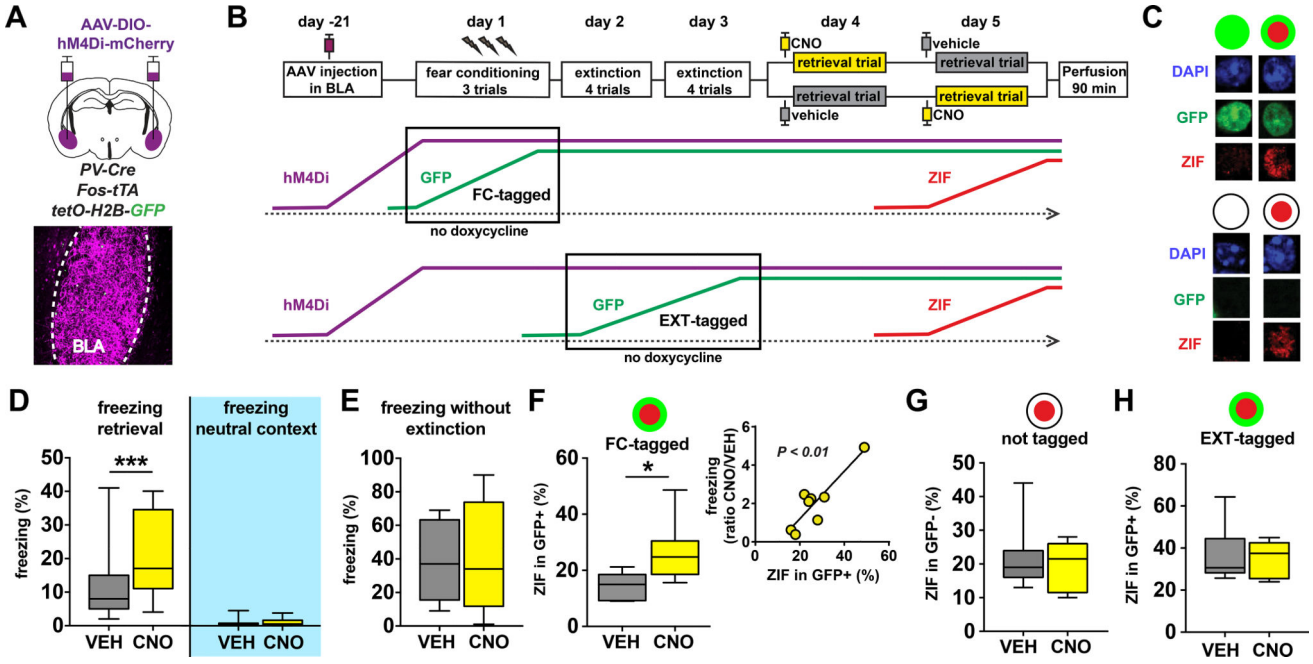

Figure 1. BLA PV interneurons selectively suppress conditioned fear behavior and neuronal ensembles following extinction

A) Bilateral infusion of AAV-Syn-DIO-hM4Di-mCherry into the BLA of PV-Cre:Fos-tTa:tetO-Hist1H2B-GFP (PV-Cre:TetTag) mice was used to selectively express DREADD receptors in BLA PV-interneurons. **B)** Mice were subjected to contextual fear conditioning, extinction, and retrieval. Cells active during the no doxycycline period are tagged with long-lasting H2B-GFP expression as a result of Fos-promoter driven tTA expression, which enables tTA protein to bind and activate the tetO-H2B-GFP transgene. This binding is prevented in the presence of doxycycline. Cells active during the retrieval trial express short-lasting Zif268 (ZIF) protein that can be detected in brains perfused 90 minutes later. PV-Cre:TetTag mice were without doxycycline chow during either the fear conditioning trials or during the extinction trials, resulting in FC-tagged and EXT-tagged neurons, respectively. For the EXT-tagged experiment, extinction and retrieval was done on days 4–7 instead of days 2–5 to enable clearance of doxycycline on days 3–4. **C)** Examples of GFP+/GFP- and ZIF+/ZIF-nuclei in the BLA of TetTag mice. **D–E)** Silencing BLA PV-interneurons results in a selective increase in conditioned fear following extinction. Mice injected with CNO 30 min prior to retrieval displayed increased freezing in the conditioned context (**D left**, Wilcoxon matched-pairs: $W = 338$, $P < 0.0001$, $n = 29$ mice), but not in a neutral context that was never paired with footshock (**D right**, Wilcoxon matched-pairs: $W = 9$, $P = 0.6563$, $n = 9$ mice). Mice injected with CNO 30 min prior to a retrieval trial that was not preceded by extinction trials did not show altered freezing levels in the conditioned context (**E**, Wilcoxon matched-pairs: $W = -1$, $P = 0.9824$, $n = 12$ mice). **F–H)** Silencing PV interneurons selectively disinhibits the BLA fear ensemble. Injection of CNO 30 min prior to retrieval leads to an increase in the percentage of ZIF+ cells among the FC-tagged GFP+ neurons (**F left panel**, unpaired t-test: $t(13) = 2.86$, $P = 0.0134$, VEH $n = 7$ mice, CNO $n = 8$ mice). The percentage reactivated GFP+ cells correlated with the behavioral effect of CNO (**F right panel**, linear regression: $F(1,6) = 23.42$, $P = 0.0029$, $n = 8$ mice). Injection of CNO

had no effect on GFP− neurons that were not tagged during fear conditioning (**G**, Mann-Whitney test: U = 27, P = 0.9333, VEH n = 7 mice, CNO n = 8 mice), or on EXT-tagged GFP+ neurons (**H**, unpaired t-test: t(10) = 0.4142, P = 0.6875, VEH n = 7, CNO n = 5). All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

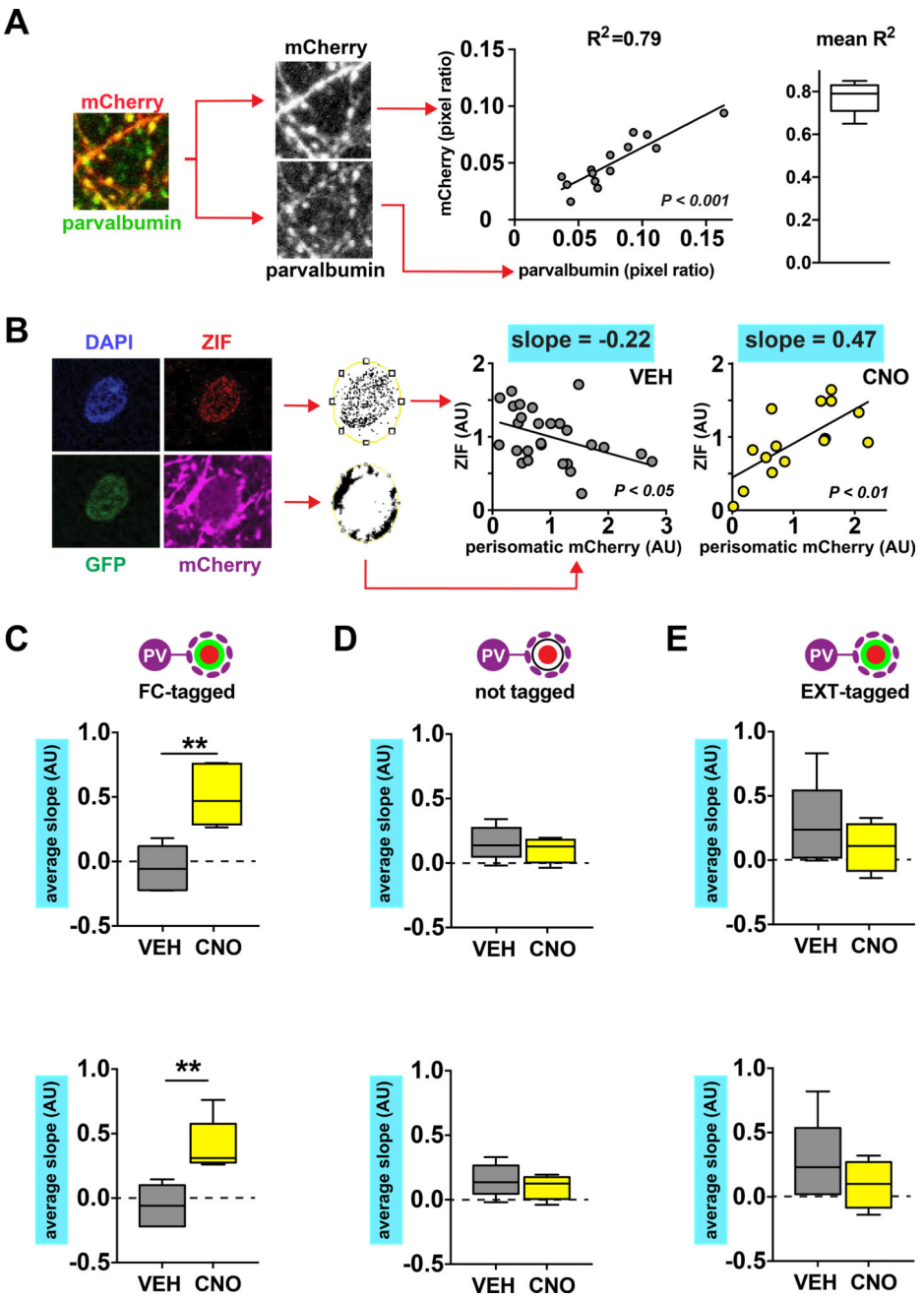

Figure 2. Structured perisomatic inhibition selectively silences BLA fear neurons during post-extinction retrieval

A) Left: example of confocal image showing overlap of mCherry and parvalbumin immunofluorescence. Middle: example data from one mouse demonstrating high degree of correlation between the local presence of perisomatic mCherry and parvalbumin (linear regression: $F(1,12) = 46.16$, $P < 0.0001$, $n = 14$ neurons). Right: mean R^2 of 5 mice analyzed as in middle panel. **B)** Left: example GFP+ZIF+ neuron demonstrating detection of perisomatic mCherry signal. Middle: example data from a VEH injected mouse with a negative correlation between perisomatic mCherry and ZIF signal (linear regression: $F(1,25) = 5.106$, $P = 0.0328$, $n = 27$ neurons). Right: example data from a CNO injected mouse with

a positive correlation between perisomatic mCherry and ZIF signal (linear regression: $F(1,13) = 11.83$, $P = 0.0044$, n = 15 neurons).**C)** CNO injection 30 min prior to post-extinction retrieval results in a positive relationship between perisomatic mCherry and ZIF expression in FC-tagged GFP+ neurons (unpaired t-test: $t(8) = 4.277$, $P = 0.0027$, VEH n = 5 mice, CNO n = 5 mice). **D–E)** CNO has no effect on the relationship between perisomatic mCherry and ZIF expression in GFP− neurons that were not tagged during fear conditioning (**D**, unpaired t-test: $t(9) = 0.798$, $P = 0.4454$, VEH n = 5 mice, CNO n = 6 mice), or in EXT-tagged GFP+ neurons (**E**, unpaired t-test: $t(8) = 1.06$, $P = 0.3199$, VEH n = 6 mice, CNO n = 4 mice). All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

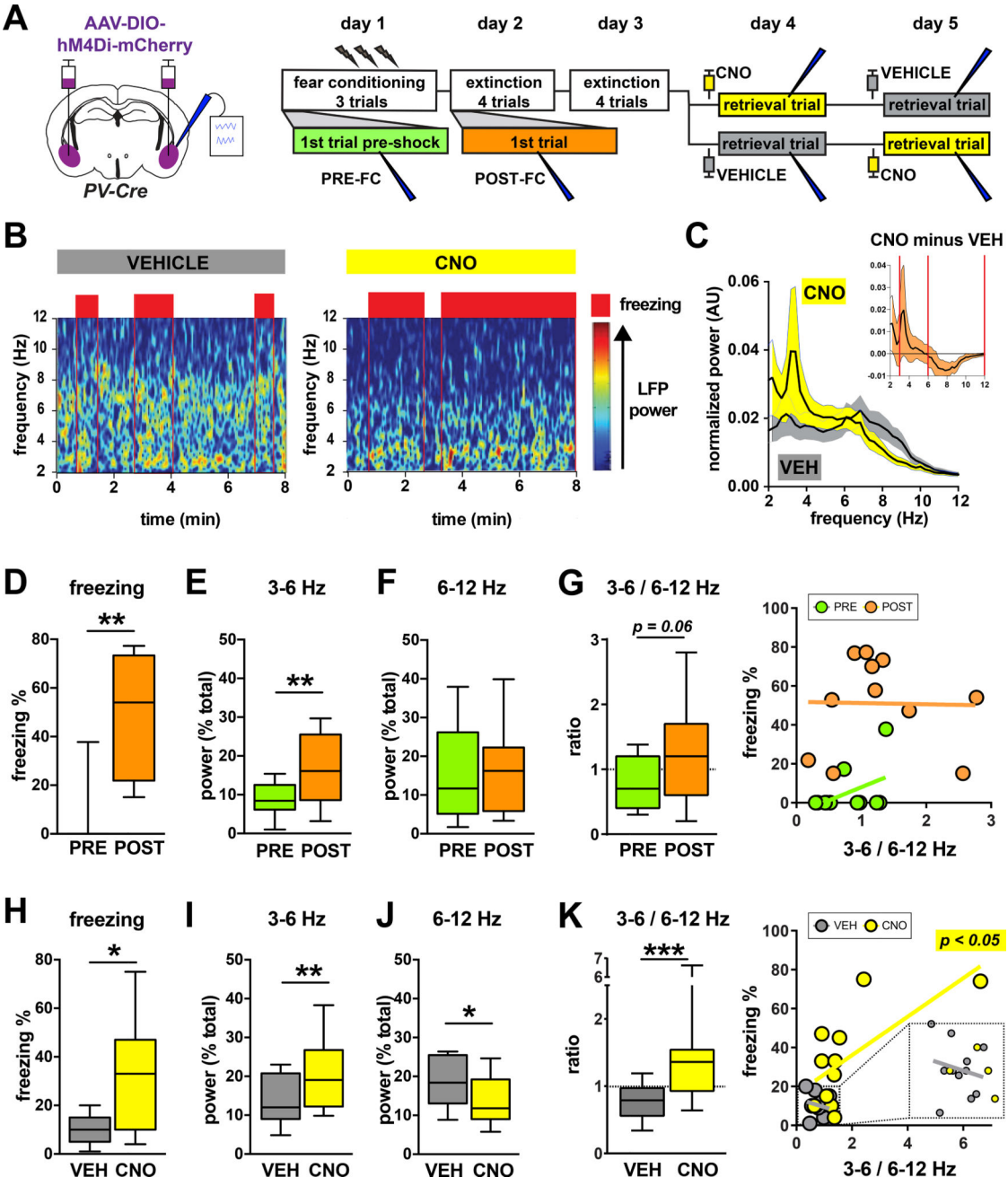

Figure 3. BLA PV-interneurons control the balance between two functionally opposed low frequency oscillations

A) Experimental design: mice were infused bilaterally with AAV-Syn-DIO-hM4Di-mCherry and simultaneously implanted with recording electrodes in BLA. Mice were then subjected to contextual fear conditioning, extinction, and retrieval. **B)** Example of full trial spectrograms from a single animal demonstrating a 3–6 Hz oscillation during freezing (red boxes: periods of >50% freezing per bin), as well as a shift towards increased 3–6 Hz power compared to 6–12 Hz power caused by silencing BLA PV interneurons. **C)** Averaged normalized 2–12 Hz power spectra from CNO and VEH retrieval trials. INSET: Running

difference between power spectra from CNO trials and VEH trials, calculated by subtracting averaged VEH trial spectra from CNO trial spectra (n = 11; red vertical lines mark 3–6 Hz and 6–12 Hz intervals used for quantification). **D–G)** Fear conditioning leads to increased freezing in the conditioned context (**D**, Wilcoxon matched-pairs: W = 64, P = 0.0020, n = 11 mice), increased 3–6 Hz power (**E**, paired t-test: t(10) = 3.481, P = 0.0059, n = 11 mice), no change in 6–12 Hz power (**F**, paired t-test, t(10) = 0.3747, P = 0.7157, n = 11 mice), and caused a trend for an increased 3–6/6–12 Hz power ratio (**G left**, paired t-test: t(10) = 2.074, P = 0.0648). The 3–6/6–12 Hz power ratio does not correlate with freezing before (**G right PRE**, linear regression: F(1,9) = 2.299, P = 0.1638, n = 11 mice), or after fear conditioning (**G right POST**, linear regression: F(1,9) = 0.003989, P = 0.9510, n = 11 mice). **H–K)** Silencing PV-interneurons during post-extinction retrieval leads to increased freezing in the conditioned context (**H**, paired t-test: t(10) = 2.867, P = 0.0167, n = 11 mice), increased 3–6 Hz power (**I**, Wilcoxon matched-pairs: W = 56, P = 0.0098, n = 11 mice), decreased 6–12 Hz power (**J**, paired t-test: t(10) = 2.427, P = 0.0356, n = 11 mice), and an increased 3–6/6–12 Hz power ratio (**K left**, Wilcoxon matched-pairs: W = 66, P = 0.0010, n = 11 mice). The 3–6/6–12 Hz power ratio correlates with freezing in CNO injected mice (**K right CNO**, linear regression: F(1,9) = 7.423, P = 0.0234, n = 11 mice), but not in VEH injected mice (**K right VEH**, linear regression: F(1,9) = 0.3594, P = 0.5636, n = 11 mice). All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

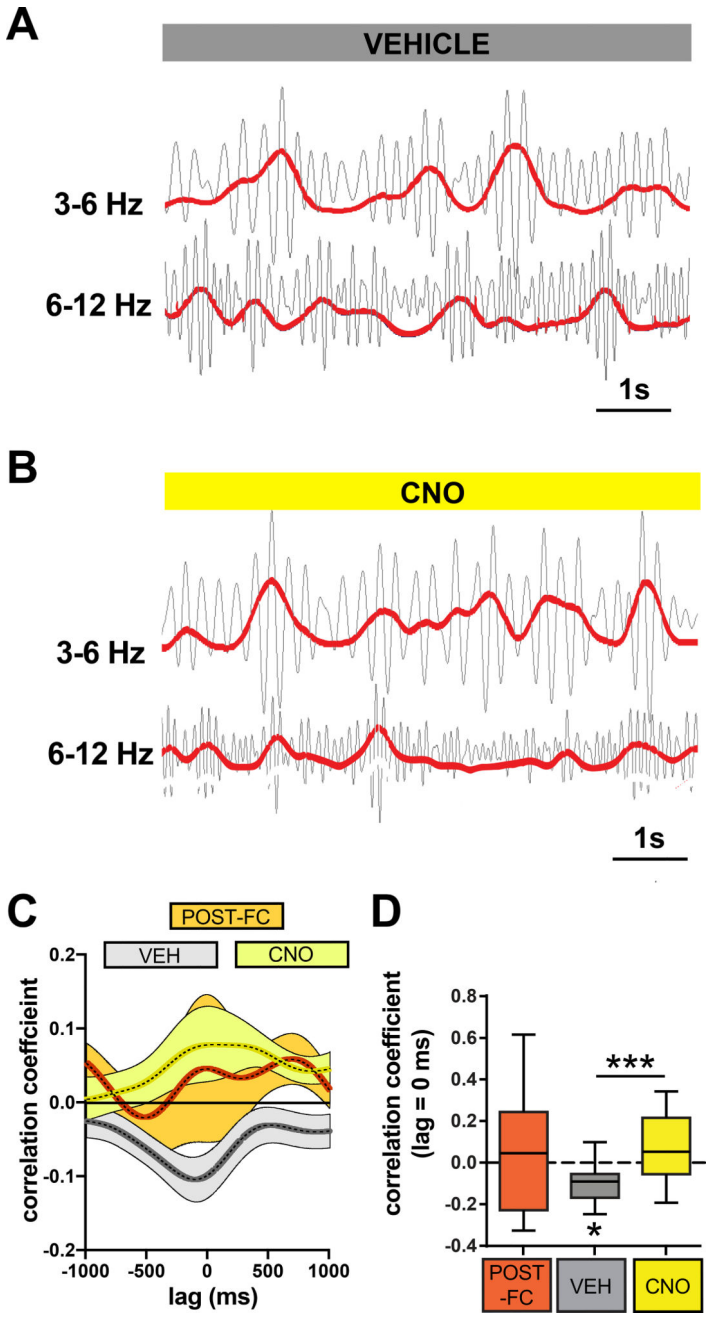

Figure 4. BLA PV-interneurons enable competition between 3–6 Hz and 6–12 Hz oscillations

A–B) Examples of 3–6 Hz and 6–12 Hz filtered traces indicate a negative correlation between the instantaneous powers of these two bands in a vehicle-injected mouse, and a positive correlation in a CNO-injected mouse. **C)** Averaged cross-correlograms were calculated for POST-FC, VEH, and CNO conditions (see experimental design in Fig. 3A). **D)** Quantification of the correlation coefficient at 0 lag revealed a significant negative correlation between the 3–6 Hz and 6–12 Hz bands after extinction in VEH injected mice (VEH One sample t-test with theoretical mean = 0: t(10) = 3.131, P = 0.0107, n = 11 mice), which was eliminated by CNO injection (VEH versus CNO Paired t-test: t(10) = 4.746, P =

0.0008, n = 11 mice). There was no significant negative or positive correlation before extinction or following CNO injection (POST-FC One sample t-test with theoretical mean = 0: t(8) = 0.4696, P = 0.6512, n = 9 mice; CNO One sample t-test with theoretical mean = 0: t(10) = 1.636, P = 0.1329, n = 11 mice). Box plot graph shows median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

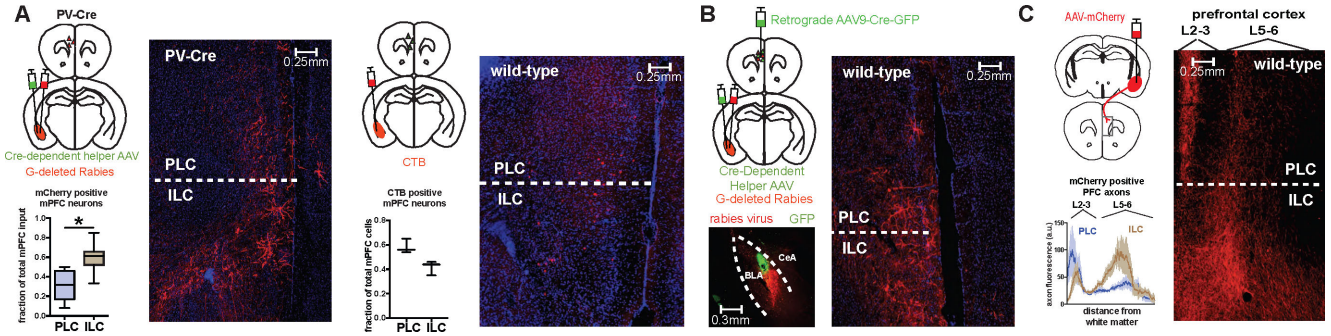

Figure 5. BLA PV-interneurons participate in a reciprocal BLA-mPFC circuit

A) Left: to trace monosynaptic inputs on BLA PV-interneurons, helper virus (AAV9-Ef1α-FLEX-GTB) and mCherry-expressing rabies virus was injected into BLA of PV-Cre mice. The example image shows mCherry-positive neurons in mPFC that synapse onto BLA PV-interneurons, with the graph showing that a larger fraction of mCherry-positive neurons was located in the ILC versus PLC (**A left**, paired t-test: t(7) = 2.8, P = 0.0265, n = 8 mice). Right: to trace all inputs to the BLA, CTB was injected into the BLA of wildtype mice. The example image shows CTB-positive neurons in the mPFC that project to the BLA, with the graph showing similar fractions of CTB-positive neurons located in the PLC and ILC (**A right**, paired t-test: t(2) = 2.463, P = 0.1328, n = 3 mice). **B)** To selectively label monosynaptic inputs onto BLA→mPFC projecting neurons, retrograde AAV9-Cre-GFP was injected into mPFC, and helper virus and mCherry-expressing rabies virus injected into BLA three weeks later. Bottom left: example image of injection site of mCherry-expressing rabies virus (red) in the BLA and green nuclei from retrograde AAV9-Cre-GFP virus injected into mPFC. Right: example image demonstrating mCherry-positive neurons in mPFC that synapse onto BLA→mPFC projecting neurons. **C)** To label BLA→mPFC projections, AAV-Cre-GFP was injected into mPFC, and AAV-Syn-DIO-hM4Di-mCherry was injected into BLA. Right and bottom left: representative image and quantification showing dense innervation of superficial PL layers (L2–3) and deep IL layers (L5–6) by BLA projection neurons (n = 4). All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

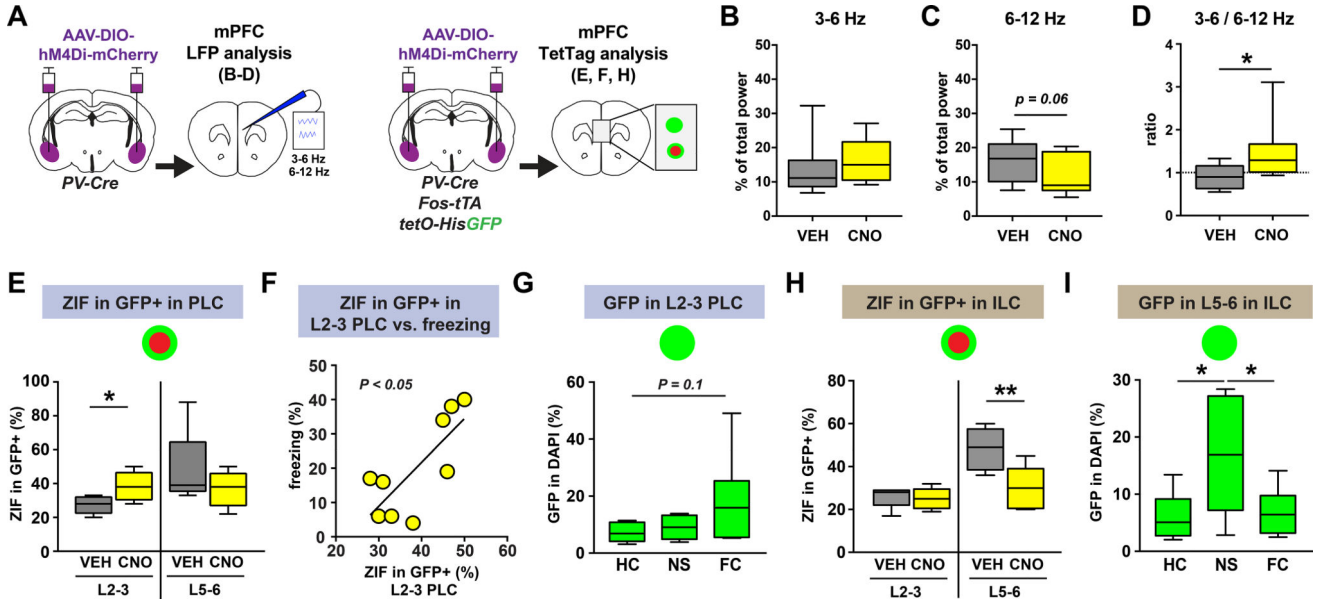

Figure 6. Silencing BLA PV-interneurons shifts mPFC ensemble dynamics

A) Chemogenetic silencing of BLA PV-interneurons was combined with either *in vivo* LFP recording in the mPFC using the same group of mice as in figure 3, or with TetTag analysis of the mPFC using the same group of mice as in figure 1. **B–D)** Silencing PV-interneurons during post-extinction retrieval increased the 3–6/6–12 Hz power ratio in the mPFC (**D**, Wilcoxon matched-pairs: W = 39, P = 0.0195, n = 9 mice), without significantly changing the 3–6 Hz power (**B**, Wilcoxon matched-pairs: W = 23, P = 0.2031, n = 9 mice), or the 6–12 Hz power (**C**, paired t-test, t(8) = 2.201, P = 0.0589, n = 9 mice). **E)** Silencing BLA PV interneurons during post-extinction retrieval increased the reactivation of superficial PLC GFP+ neurons (L2–3, unpaired t-test, t(12) = 2.689, P = 0.0197, VEH n = 5 mice, CNO n = 9 mice), but had no effect on deep PLC GFP+ neurons (L5–6, Mann-Whitney test: U = 17, P = 0.4935, VEH n = 5 mice, CNO n = 9 mice). **F)** The percentage of reactivated GFP+ neurons in the superficial PLC correlated with freezing behavior (linear regression: F(1,7) = 9.977, P = 0.0160, n = 9 mice). **G)** In a separate experiment, fear conditioning caused a trend for increased GFP+ tagged neurons in the superficial PLC as compared to mice left in the homecage (one-way ANOVA: F(2,17)=2.684, P = 0.0970, homecage/HC n = 8 mice, no-shock/NS n = 5 mice, fear conditioned/FC n = 7 mice; Tukey's multiple comparisons test HC versus FC: q(17) = 3.139, P = 0.0964). **H)** Silencing BLA PV interneurons during post-extinction retrieval decreased the reactivation of deep ILC GFP+ neurons (L5–6, unpaired t-test: t(12) = 3.336, P = 0.0059, VEH n = 5 mice, CNO n = 9 mice), but had no effect on superficial GFP+ neurons (L2–3, Mann-Whitney test: U = 20, P = 0.7962, VEH n = 5 mice, CNO n = 9 mice). **I)** In a separate experiment, deep layer ILC neurons were preferentially tagged in animals exposed to a novel context without shock, as compared to homecage and fear conditioned animals (one-way ANOVA: F(2,19) = 5.282, P = 0.0150, homecage/HC n = 8 mice, no-shock/NS n = 6 mice, fear conditioned/FC n = 8 mice; Tukey's multiple comparisons test HC versus NS: q(19) = 4.209, P = 0.0202; Tukey's multiple comparisons test NS versus FC: q(19) = 3.921, P = 0.0311). All box plot graphs show median (line inside

box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).

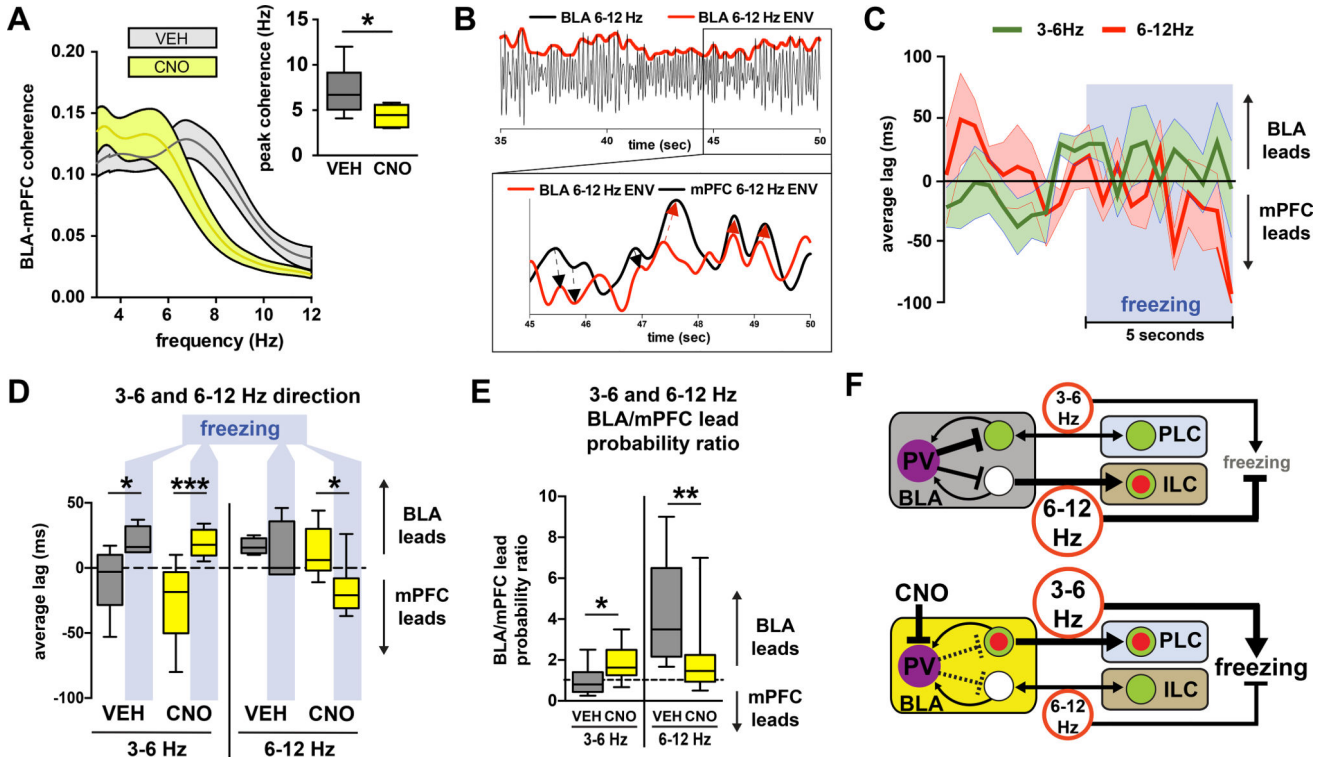

Figure 7. BLA PV-interneurons control directionality of 3–6 and 6–12 Hz oscillations following extinction

A) Silencing BLA PV-interneurons during post-extinction retrieval shifts BLA-mPFC peak coherence towards the 3–6 Hz range (paired t-test: t(8) = 2.486, P = 0.0378, n = 9 mice). **B)** Example of a 6–12 Hz filtered trace demonstrating the extraction of instantaneous power envelopes (top) which can be cross-correlated (bottom) to determine leading (black arrows) or lagging (red arrows) relationships between two signals. **C)** Averaged lags from instantaneous power cross-correlations for both 3–6 Hz and 6–12 Hz bands during the 5 seconds preceding and the 5 seconds following the onset of freezing during retrieval trials (VEH n= 6 mice, CNO n = 8 mice, 1 freezing event per mouse per trial). **D)** The first 5 seconds following the onset of freezing is characterized by a shift towards a BLA lead for 3–6 Hz in both VEH and CNO groups (repeated measures two-way ANOVA: freezing F(1,11) = 29.12, P = 0.0002, CNO F(1,11) = 1.013, P = 0.3357, freezing×CNO F(1,11) = 1.411, P = 0.2598, VEH n = 5 mice, CNO n = 8 mice; Sidak's multiple comparisons test: pre-freeze vs freeze for VEH t(11) = 2.682, P = 0.0422; Sidak's multiple comparisons test: pre-freeze vs freeze for CNO t(11) = 5.308, P = 0.0005), and a shift towards a mPFC lead for 6–12 Hz in the CNO group (repeated measures two-way ANOVA: freezing F(1,9) = 5.487, P = 0.0439, CNO F(1,9) = 2.263, P = 0.1668, freezing×CNO F(1,9) = 2.355, P = 0.1592, VEH n = 4 mice, CNO n = 7 mice; Sidak's multiple comparisons test: pre-freeze vs freeze for VEH t(9) = 0.5063, P = 0.8593; Sidak's multiple comparisons test: pre-freeze vs freeze for CNO t(9) = 3.215, P = 0.0210). **E)** Silencing PV interneurons during post-extinction retrieval increased the probability of BLA leading mPFC in the 3–6 Hz band (paired t-test, t(9) = 2.303, P = 0.0468, n = 10 mice), while decreasing the probability that it will lead in the 6–12 Hz band (Wilcoxon matched-pairs: W = −45, P = 0.0039, n = 9 mice). **F)** Model based on our data

and previous studies. Top: after extinction learning, BLA PV-interneurons suppress conditioned freezing and fear ensemble activation by enabling a 6–12 Hz oscillation to out-compete a 3–6 Hz oscillation throughout the BLA-mPFC circuit. Bottom: chemogenetic silencing of BLA PV-interneurons causes dysfunctional competition between the 6–12 and 3–6 Hz oscillations, which leads to an increase in 3–6 Hz power and BLA→mPFC directionality, an increase in the activation of fear ensembles in BLA and PLC, and an increase in conditioned freezing. All box plot graphs show median (line inside box), 25% and 75% percentiles (box edges), and minimum and maximum values (error bars).