Selective control of fear expression by optogenetic manipulation of infralimbic cortex after extinction

Hyung-Su Kim[1,2], Hye-Yeon Cho[1], George J. Augustine[2-4], Jin-Hee Han[1]

[1]Department of Biological Sciences, KAIST Institute for the BioCentury (KIB), Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Korea

[2]Center for Functional Connectomics, Korea Institute of Science and Technology, Hwaran gno 14-gil 5, Seongbuk-gu, Seoul 136-791, Korea

[3]Lee Kong Chian School of Medicine, Nanyang Technological University, 50 Nanyang Drive, Research Techno Plaza, Singapore 637553, Singapore

[4]Institute of Molecular and Cell Biology, A*STAR, Proteos, 61 Biopolis Drive, Singapore 138673, Singapore

Running title: Retrieval of fear extinction by infralimbic cortex

Address for correspondence: Jin-Hee Han

Department of Biological Sciences, Korea Advanced Institute of Science and Technology

291 Daehak-ro, Yuseong-gu, Daejeon, Korea 305-701

Tel: 82-42-350-2649

E-mail: han.jinhee@kaist.ac.kr

Abstract

Evidence from rodent and human studies has identified the ventromedial prefrontal cortex, specifically the infralimbic cortex (IL), as a critical brain structure in the extinction of conditioned fear. However, how IL activity controls fear expression at the time of extinction memory retrieval is unclear and controversial. To address this issue, we used optogenetics to precisely manipulate the activity of genetically targeted cells and to examine the real-time contribution of IL activity to expression of auditory conditioned fear extinction in mice. We found that inactivation of infralimbic cortex, but not prelimbic cortex, impaired extinction retrieval. Conversely, photostimulation of IL excitatory neurons robustly enhanced the inhibition of fear expression after extinction, but not before extinction. Moreover, this effect was specific to the conditioned stimulus (CS): IL activity had no effect on expression of fear in response to the context after auditory fear extinction. Thus, in contrast to the expectation from a generally held view, artificial activation of IL produced no significant effect on expression of non-extinguished conditioned fear. Therefore, our data provide compelling evidence that IL activity is critical for expression of fear extinction and establish a causal role for IL activity in controlling fear expression in a CS-specific manner after extinction.

Introduction

Defining the precise role of neural circuits involved in the regulation of fear expression is critical for understanding and treating fear and anxiety disorders such as post-traumatic stress disorder (PTSD). Learned fear has been extensively studied using classical fear conditioning paradigm in rodents, particularly auditory fear conditioning, in which a tone (conditioned stimulus [CS]) is associated with a foot-shock (unconditioned stimulus [US]; Fanselow and Gale, 2003; Kapp *et al*, 1979; LeDoux, 2000). After conditioning, the tone alone can induce freezing—an expression of fear—in conditioned animals. The conditioned fear response to a tone CS can be gradually weakened by repeated exposure to unreinforced tone without the US, a process called extinction (Pavlov, 1927).

A number of studies employing various techniques have specifically implicated the infralimbic cortex (IL) within the vmPFC area in controlling fear expression after extinction in rodents and humans (Milad and Quirk, 2002, 2012; Milad *et al*, 2006; Phelps *et al*, 2004; Santini *et al*, 2008). Lesions targeting the vmPFC impair extinction learning and expression (Morgan *et al*, 1993; Quirk *et al*, 2000). Specifically, IL lesions within the vmPFC in rats extinguish the freezing response to tone within an extinction session, but excessive freezing spontaneously recovers the following day, suggesting failure of extinction recall (Lebron *et al*, 2004; Quirk *et al*, 2000). In vivo recordings of IL activity in freely moving rats or mice have shown increased CS-related activity and induction of synaptic plasticity in the IL following extinction training (Herry and Garcia, 2002; Hugues *et al*, 2006; Milad *et al*, 2002). In addition, pharmacological inactivation of the IL prior to extinction training results in failure of extinction recall (Hugues *et al*, 2006; Laurent and Westbrook, 2009; Morgan and LeDoux,

1995; Morgan *et al*, 1993). It has also been suggested that the synaptic connection between the IL and amygdala is crucial for fear expression after extinction (Cho *et al*, 2013; Likhtik *et al*, 2008; Myers and Davis, 2007; Senn et al., 2014). In human studies, vmPFC activity apparently is correlated with success of extinction expression (Milad *et al*, 2007; Phelps *et al*, 2004). These findings support the hypothesis that increased activity in the IL mediates inhibition of fear responses to the tone CS after extinction.

However, experimental analyses of the involvement of the IL in extinction recall via lesion or inactivation approaches have yielded inconsistent results. Some previous studies have reported that lesions to the vmPFC, or specifically to the IL, fail to impair extinction memory recall (Farinelli *et al*, 2006; Garcia *et al*, 2006; Gewirtz, 1997; Morgan *et al*, 2003). Further, inactivation of the vmPFC immediately before extinction retrieval - by infusion of the sodium channel blocker, tetrodotoxin - facilitated rather than impaired extinction (Sierra-Mercado *et al*, 2006). Moreover, CS-evoked firing in the IL reportedly is greater in rats that fail to acquire extinction (Chang *et al*, 2010). In a very recent study, optogenetic silencing of IL activity during the retrieval of extinction of auditory conditioned fear had no effect (Do-Monte *et al*, 2015). The reason for these divergent results is unclear, but the differences probably arise from different experimental conditions, in particular the spatiotemporal precision of the techniques used to manipulate neuronal activity. For instance, the prelimbic cortex (PL), which is located immediately above the IL, is known to have opposite influences on fear expression and thus precise and very careful manipulation must be applied to clearly determine the role of each division of the mPFC (Sierra-Mercado *et al*, 2011; Vidal-Gonzalez *et al*, 2006). Most commonly used manipulations of neural activity, such as lesions, drug-mediated inactivation, and electrical stimulation, provide relatively poor spatiotemporal resolution. These techniques also suffer from potential non-specific side effects, such as

functional compensation by other brain structures (lesions), inaccurate control of diffusion (drugs), and unwanted, non-specific stimulation of axon fibers crossing target brain regions (electrical stimulation). Therefore, despite much effort, it remains unclear whether and how vmPFC, particularly IL, controls expression of fear extinction.

To address the role of IL in fear extinction, we employed optogenetics to manipulate the activity of genetically targeted cells in the IL of mice in a temporally and spatially precise manner. This allowed us to investigate the behavioral effects of bidirectional manipulation of neuronal activity in the IL specifically at the time of retrieval on the expression of conditioned fear before and after extinction.

Materials and Methods

Subjects. Hybrid male 129/C57B1/6 mice (2–3 months old, 23–35 g) were used for all experiments. Mice were group housed and maintained in a 12 h light/dark cycle at a consistent temperature (22 ± 2 °C) at 40–60% humidity. Mice were provided food and water *ad libitum* throughout experiments. All procedures were approved by the Animal Ethics Committee at each institution where the experiments were performed.

Surgery. AAV virus for brain surgery was prepared as described previously (Kwon *et al*, 2014). Five AAV viral vector constructs were used in this study (1–5 × 10^{11} vg mL^{-1}): AAV-hSyn-eNpHR3.0-EYFP, AAV-hSyn-EGFP, AAV-CaMKIIα-eNpHR3.0-EYFP, AAV-CaMKIIα-ChR2-Venus, and AAV-CaMKIIα-EGFP. Mice were deeply anesthetized with pentobarbital (83 mg kg^{-1} of body weight) by intraperitoneal injection and placed in a stereotaxic frame. A glass micropipette was loaded with the virus and positioned in the IL (anterior-posterior [AP] = +1.65 mm, mediolateral [ML] = ±0.35 mm, dorsoventral [DV] = –3.05 mm) or PL (AP = +1.7 mm, ML = –0.35 mm, DV = –2 mm). AAV solutions were injected at a rate of 0.1 μl min^{-1} for 5 min (total, 0.5 μl) in the IL or 7 min (total, 0.7 μl) in the PL. The micropipette was left in place for an additional 10 min to allow virus diffusion and withdrawn. Subsequent experiments were conducted 2–3 weeks later to ensure sufficient time

for expression of opsins in the targeted region. Afterwards, a guide cannula (Plastics One, VA, USA) was placed into the IL (AP = +1.6 mm, ML = ±1.3 mm, DV = –2.3 mm), 15 degrees angled to the midline to avoid damage to the vmPFC or PL (AP = +1.7 mm, ML = –0.35 mm, DV = –1 mm) and fixed with dental cement.

In vitro patch-clamp recordings in brain slices. Three weeks after AAV viral injection into IL, mice were deeply anesthetized with halothane and decapitated, and their brains were quickly removed. Coronal slices (350-μm thick) were prepared in ice-cold sucrose solution and incubated in artificial cerebrospinal fluid (ACSF) at 36 °C for 30 min and at room temperature for at least 30 min prior to use. All of the electrophysiological recordings were performed with continuous perfusion of 32 °C oxygenated ACSF. The patch electrode (4–12 MΩ) was filled with an internal solution. Detailed information of composition of solutions and equipment for data acquisition used in this experiment are available in **Supplementary Materials and Methods**. A mercury arc lamp was used to activate ChR2 (470–495 nm) and eNpHR3.0 (530–550 nm). Light stimuli were delivered through a water-immersion objective (60×). Photocurrents in eNpHR3.0- or ChR2-expressing IL neurons were directly measured in voltage-clamp mode at a holding potential of -50 mV in response to different intensities of light delivered for 2 s. To obtain reliable action potentials, 200 pA currents were injected. To determine the fidelity of ChR2-evoked spikes, 4-ms light pulses of different frequencies (5, 10, 20, 30, and 40 Hz) were delivered at 4.6 mW mm^{-2} for 1 s. Pulse duration was controlled by an electronic shutter (Uniblitz VS25, Vincent, NY, USA).

Behavioral procedure. Auditory fear conditioning and extinction: Seven days after recovery from cannula implantation, all mice were handled and habituated by inserting an optic fiber

prior to behavioral experiments. Mice were submitted to three phases of behavioral procedures for extinction experiment: fear conditioning, extinction training, and extinction retrieval. For auditory fear conditioning (day 1), mice were placed in a conditioning chamber (Coulbourn Instruments, PA, USA) and allowed to explore for 2 min (pre-CS). Subsequently, a CS tone (2.8 kHz, 85 dB, 30 s) was paired with an aversive foot-shock US (0.5 mA, 2 s) three times with a variable, 100–120-s inter-stimulus interval (ISI). Each tone presentation was co-terminated with an electrical foot-shock. Mice were kept in the chamber for an additional 1 min after delivery of the last shock for monitoring of post-shock behavior. Twenty-four hours later (day 2), mice connected to the optic fiber (200 μm core; Thorlabs, NJ, USA) were placed in a novel context-shifted chamber which consisted of a plastic door covered with a vertical stripes, a white acrylic floor and semicircular cylindrical wall for extinction training. After a 2-min exploratory period, mice received repeated presentations of a tone stimulus (40 trials with 5-s ISI, 30 s for each trial) in the absence of a foot-shock. On the next day (day 3), mice underwent the same extinction training procedure. On the test day (day 4), freezing responses were measured in 8 trials of tone presentation (30 s for each trial with 5-s ISI) after establishing a 2-min baseline freezing level. Because there was a significant positive correlation between extinction ratio across the entire training, 100*[(initial conditioned freezing – freezing during the last extinction trial block on day3)/initial conditioned freezing], and long-term extinction ratio in our conditions, the extinction ratio during training was used as an index to determine that long-term extinction occurred.

Optogenetic manipulation during extinction retrieval: On the test day, mice were placed in the same chamber as used for extinction training and their baseline level of freezing was measured freezing for 2 min (pre-CS). Immediately thereafter, light was delivered to the

target brain area during the first four trials of tone presentation (light-ON) and then was turned off for the next four trials (light-OFF). Green light (561 nm, 63.7–127.4 mW mm^{-2}) or blue light (473 nm, 6.4 mW mm^{-2}) was delivered from diode laser (CL561-050 and CL473-050, respectively, Crystalaser, NV, USA) via rotary joint. For eNpHR3.0, green light was continuously delivered (including during 4 × CS and 3 × ISIs). For ChR2, blue light (10 Hz, 20-ms pulse duration) was delivered; no light was delivered during the ISI period. The freezing values for light-ON and light-OFF conditions were determined by calculating the mean percentage of freezing from four trials each.

Photoinhibition of IL without CS: The same group of mice that received extinction retrieval tests was placed back into the same camber. Two minutes after entering the chamber, mice received the same green light as above for 135 s in the absence of tone CS presentation (see **Figure 1h**).

Optogenetic manipulation during fear memory expression: Mice were fear conditioned as described above, but with a single pairing of CS and US. Fear memory was tested in the context-shifted chamber 24 h later. After a 2-min exploratory period (pre-CS), tone was presented for 2 min to determine fear memory expression. During the first 1 min of tone presentation, either green (continuous) or blue (10 Hz, 20-ms duration) light was delivered to the target brain area (light-ON). During the next 1 min of tone presentation, freezing was measured in the absence of light illumination (light-OFF). For unilateral (20 Hz, 10-ms duration) and bilateral (10 Hz) photoactivation experiments, mice received three pairings of tone and shock presentations to avoid a floor effect. Twenty-four hours later, mice were placed in a context-shifted chamber. After 2 min of baseline recording (pre-CS), CS was presented with unilateral 20 Hz or bilateral 10 Hz light stimulation (4 trials with 5-s ISI, 30 s for each trial). Light stimulation was delivered either with or without a 0.1-s delay after tone

onset. For photoactivation during contextual fear retrieval, mice received fear conditioning and extinction training as described above (day 1–3) and on the test day (day 4), they were re-introduced the conditioned chamber for 2-min with 10Hz light stimulation during first minute. Mouse behavior was recorded using FreezeFrame software (Actimetrics, IL, USA). Freezing level during fear conditioning was determined by automatic scoring using FreezeFrame software, while freezing for all other cases was manually scored in a blinded manner by measuring the time animals spent in freezing.

Histology. At the end of behavioral experiments, all the mice were sacrificed and brain sections were prepared as described previously (Kwon *et al*, 2014). For histological verification of virus expression and fiber optic placement, fluorescence images of coronal sections were captured with a fluorescence microscope (ECLIPSE 80i, Nikon). To confirm expression in the IL region, the extent of expression was measured with scale bars (NIS Elements Software, Nikon) by reference to the Mouse Brain Atlas (Franklin and Paxinos, 1997). Only mice that showed highly restricted expression of targeted opsin in the IL region were included in data analyses. Animals that were excluded typically showed faint expression in the IL, off-target expression in the PL or dorsal tenia tecta, or physical damage in the IL caused by the guide cannula. Histological verifications were conducted in a blinded manner.

Statistical analysis. Statistical significance of data was determined using Student's *t* test or two-way repeated-measures ANOVA followed by Bonferroni *post hoc* test for multiple comparisons. A significance level of $p<0.05$ was used for all analyses. Prism (version 6.04, GraphPad Software) was used for all statistical analyses.

Results

Optogenetic inhibition of neuronal activity in the IL

We employed an adeno-associated virus (AAV) system in which expression of genes for third-generation halorhodopsin (eNpHR3.0), a light-activated chloride pump~~,~~ fused to enhanced yellow fluorescent protein (EYFP)~~,~~ or enhanced green fluorescent protein (EGFP; control) was driven by the neuron-specific human synapsin (hSyn) promoter. AAV-eNpHR3.0 was delivered into the right IL by unilateral, stereotaxic injection, followed by chronic implantation of cannulae, angled 15° toward the midline to minimize damage to the dorsal part of the mPFC (**Figure 1a**). We optimized the injection volume so that the virus-infected region covered the IL, but with little or no unwanted infection in surrounding areas. For example, virus infections were sometimes observed in the dorsal peduncular cortex (DP), particularly when expression was relatively strong. Because the DP, as part of the vmPFC, is thought to either contribute to extinction learning in a manner similar to that of the IL or to have no effect (LaLumiere *et al*, 2010; Peters *et al*, 2009; Van den Oever *et al*, 2013;

Willcocks and McNally, 2013), we also included these animals in our data analysis. In addition, after behavioral experiments we checked the size of the virus-infected areas in each mouse, and included in our final analysis only mice that displayed highly restricted expression of eNpHR3.0 in the IL (**Figure 1a, b**). Next, we determined the proportion of glutamatergic or GABAergic neurons expressing NpHR driven by the hSyn promoter. This was done via immunohistochemistry, using antibodies against calcium/calmodulin-dependent protein kinase IIα (CaMKIIα) or glutamic acid decarboxylase67 (GAD67). The co-localization analysis showed that most cells expressing NpHR were CaMKIIα-positive excitatory neurons (69.2 ± 3.9% of NpHR-positive cells), while approximately 24% of NpHR-positive cells were GABAergic interneurons (23.8 ± 4.0%; **Supplementary Figure 1**). Thus, both excitatory and inhibitory neurons could be photoinhibited, though excitatory neurons were 3 times more abundant.

To verify the physiological effects of eNpHR3.0 on IL neuronal activity, we performed whole-cell patch-clamp recordings in brain slices. We first examined how eNpHR3.0-mediated photocurrents varied with light intensity. Illumination with 540 ± 10 nm wavelength light evoked robust hyperpolarizing currents in recorded neurons that showed a graded increased in size with increases in light intensity (**Figure 1c, d**). Next, we measured the relationship between light intensity and spike firing inhibition ratio. Under conditions in which spike firing was evoked by injection of a depolarizing current (200 pA), light at an intensity of 0.07 mW mm^{-2} inhibited spike firing by approximately 50% (46.5 ± 14.7%), whereas a 10-fold increase in light intensity (0.7 mW mm^{-2}) almost completely suppressed spike firing (87.4 ± 9.3%). The amount of inhibition was slightly increased with a further increase in light intensity, causing an inhibition of 92.0 ± 8.0% at a maximum light intensity of 7.7 mW mm^{-2} (**Figure 1e, f**). Thus, illumination evokes hyperpolarization in NpHR3.0-

expressing IL neurons and effectively inhibits spike firing in these neurons in a temporally precise manner.

Photoinhibition of IL neuronal activity impairs the expression of fear extinction

To investigate the effect of IL photoinhibition on the expression of extinction memory, we first confirmed that our protocol reliably formed long-term memory of fear extinction. Mice injected with the control EGFP virus were trained for auditory fear-conditioning and extinction. During extinction training, mice in the extinction group exhibited gradually decreased freezing by day 2 and 3 extinction sessions (**Supplementary Figure 2a**). On the test day (day 4), freezing in the extinction group (39.6%) was significantly lower than in the no-extinction group (66.4%; $p<0.5$, Student's *t* test; **Supplementary Figure 2a**). Additionally, CS-induced freezing in the extinction group (39.6%) was significantly diminished compared with initial conditioned freezing (62.5%), measured prior to extinction in the first trial block on day 2 ($p<0.01$, paired Student's *t* test); however, this was not the case for mice in the no-extinction group (66.4% and 62.5%; $p=0.7529$, Student's *t* test; **Supplementary Figure 2b**). Thus, these results confirm that the behavioral protocol efficiently induces long-term fear extinction.

Next, we tested whether IL activity is necessary for the expression of extinction memory using optogenetic inhibition. Do-Monte *et al*. (2015) recently reported that photoinhibition of excitatory neurons in IL does not affect expression of extinction memory. We therefore first targeted these excitatory neurons by expressing eNpHR3.0-EYFP or control EGFP under the CaMKIIα promoter. Mice were injected with AAV viruses into the right IL and then trained for auditory fear conditioning and extinction (**Supplementary Figure 3a, b, c**). Animals showed significant fear extinction during training (Ext1:

$F_{9,135}$=12.37, p<0.0001, Ext2: $F_{9,135}$=6.259, p<0.0001) with no significant group difference ((Ext1: $F_{1,15}$=3.083, p=0.0995, Ext2: $F_{1,15}$=3.129, p=0.0972; **Supplementary Figure 3c).** In this condition, photoinhibition of IL had no effect on expression of fear extinction (group × trial interaction, $F_{1,15}$=1.666, p=0.2163, two-way repeated-measures ANOVA; **Supplementary Figure 3d**), ~~consistent with a recent study (Do-Monte *et al*., 2015)~~. This confirms the previous findings of Do-Monte *et al*. (2015)

We next asked whether a more general silencing of both glutamatergic and GABAergic neurons ~~may have an e~~could affect ~~on~~ expression of extinction. For this purpose, we expressed eNpHR3.0-EYFP in the IL neurons using neuron-specific hSyn promoter. Virus-injected mice were subjected to fear conditioning, extinction training and testing (**Figure 1g**). Notably, there were no significant differences between EGFP control and NpHR groups in initial conditioned freezing (p=0.49, Student's t test) or in the extinction ratio during training (**Figure 1h**). Two-way repeated-measures analysis of variance (ANOVA) revealed a significant effect of trial block (Ext1: $F_{9,144}$=6.139, p<0.0001, Ext2: $F_{9,144}$=3.323, p=0.001) with no significant group difference (Ext1: $F_{1,16}$=0.01911, p=0.8918, Ext2: $F_{1,16}$=0.04034, p=0.8434; **Figure 1h**). During extinction retrieval testing, pre-CS freezing levels did not differ between groups (14% and 17.1% for EGFP and eNpHR3.0, respectively; p=0.586, Student's t test; **Figure 1i**). However, upon presentation of a tone combined with green light (561 nm) illumination, levels of freezing in eNpHR3.0 mice were significantly higher than those in control mice (**Figure 1i;** p<0.05, Student's t test), indicating that the expression of fear extinction memory was impaired by photoinhibition of IL activity at the time of retrieval. This impairment was reversible and specific: freezing of eNpHR3.0 mice returned to control levels in the absence of photoinhibition during the light-OFF phase (35% and 36.5% for EGFP and eNpHR3.0, respectively; p=0.7802, Student's t test; **Figure 1i**).

Two-way repeated-measures ANOVA confirmed a significant group difference in freezing during the light-ON period (group × trial interaction, $F_{1,16}$=8.818, p<0.01; Bonferroni *post hoc* test, p<0.05). To further confirm the specificity of the behavioral effects, we also analyzed the data from animals (n = 5) that were excluded due to no or low viral infection, or off-target expression. These mice showed a trend toward a low level of freezing during the light-ON period (19.8 ± 7.7%), similar to the control group, further supporting that impairment of extinction expression was specific to optogenetic inhibition of IL neuron activity. Hence, these results indicate that impairment of extinction retrieval was highly specific to silencing activity of neurons targeted by hSyn promoter, meaning both glutamatergic and GABAergic neurons.

Although unlikely, it is possible that IL photoinhibition may simply induce freezing even in the absence of CS. To test this possibility, we re-introduced the same mice from each group into the test chamber 1 day after extinction retrieval tests. The green light was again delivered to IL, but this time without the CS tone (**Figure 1j**). We found no significant induction of freezing compared to baseline in either eNpHR3.0 or control groups under these conditions (group × trial interaction, $F_{1,18}$=0.6643, p=0.4257, two-way repeated-measures ANOVA; **Figure 1j**), thus excluding the possibility that photoinhibition of neuronal activity in the IL directly induced freezing. Taken together, these results indicate that ~~IL~~ activity of both excitatory and inhibitory neurons in IL is necessary for the expression of fear extinction memory.

No effect of IL photoinhibition on conditioned fear before extinction

It has been suggested that the IL is engaged for fear inhibition after extinction (Sotres-Bayon and Quirk, 2010). We tested this idea by examining the effects of IL photoinhibition on fear

memory expression before extinction (**Figure 1k**). We observed no significant differences in freezing between eNpHR3.0 and control groups during light-ON or light-OFF periods (group × trial interaction, $F_{1,9}$=0.3285, p=0.5806), indicating no significant effect of IL inactivation on the expression of conditioned fear that did not undergo extinction (**Figure 1k**). Thus, IL activity is specifically engaged for the expression of fear extinction, but not of fear memory *per se.* This is consistent with previous findings (Sierra-Mercado et al., 2011) and also shows that IL photoinhibition does not simply enhance CS-evoked freezing.

Inhibition of PL activity does not affect extinction memory retrieval

A recent study reported that neuronal activity in the PL area, in addition to the IL, is increased during extinction memory retrieval (Chang *et al*, 2010). This suggests a possible role for the PL in extinction recall. In order to test this possibility, we injected AAV-eNpHR3.0 or AAV-EGFP viral vector specifically into the PL area of the vmPFC and mice were then trained for auditory fear conditioning followed by extinction training as before (**Figure 2a–c**). During extinction retrieval tests, mice in both groups displayed normal expression of fear extinction with no significant difference in freezing between groups. Two-way repeated-measures ANOVA showed a significant effect of interaction (group × trial interaction, $F_{1,16}$=7.195**,** p=0.0164) but no main effect of group ($F_{1,16}$=0.01264, p=0.9119). Bonferroni *post hoc* tests revealed no significant difference in freezing between eNpHR3.0 and control groups under light-ON (p=0.406) or light-OFF (p=0.2218) conditions during the test (**Figure 2d**). These results indicate that PL activity is not critical for the expression of extinction memory and further validate the selective role of the IL in extinction memory retrieval.

Optogenetic activation of excitatory neurons in the IL

Next, we asked whether IL activation is sufficient to inhibit fear expression. For this purpose we photostimulated IL activity by expressing the light-activated cation channel, channelrhodopsin-2 (ChR2), in excitatory neurons. To selectively target excitatory neurons in the IL, we employed an AAV-ChR2-Venus viral vector with the CaMKIIα promoter (Benson *et al*, 1992; Liu and Jones, 1996). Immunohistochemical analyses confirmed selective expression of ChR2 in CaMKIIα-positive neurons in the IL (**Figure 3a, b**). To examine the ability of ChR2 to photostimulate IL neurons, we performed whole-cell patch clamp recordings in brain slices. Large inward photocurrents were induced in ChR2-expressing neurons exposed to blue (470–495 nm) light. Increasing light intensity increased photocurrent amplitude, with a maximal light intensity of 4.6 mW mm^{-2} evoking peak photocurrents of 273 ± 56 pA (**Figure 3c, d**). In current clamp recordings, illumination evoked action potentials in IL neurons (**Figure 3e**). Action potentials were produced by light with high probability at relatively low frequencies (5 or 10 Hz), while the probability of photostimulation gradually declined at higher frequencies up to 40 Hz (**Figure 3f**). Thus, ChR2-mediated photostimulation elicits high-fidelity activation of IL neurons.

Photostimulation of IL excitatory neurons enhances expression of fear extinction

If IL activation controls the inhibition of fear expression, then increasing IL activity should enhance the expression of fear extinction. To test this idea, we trained mice in which virus expressing either ChR2-Venus or EGFP under the CaMKII promoter was injected into ~~in~~ the right IL. These mice were then subjected to auditory fear conditioning and subsequent extinction, as before (**Figure 3g**). Importantly, animals in control and ChR2 groups both showed similar levels of conditioned freezing (p=0.6014, Student's t test) and extinction during training (**Figure 3h**). Two-way repeated-measures ANOVA revealed a significant

effect of trial block (Ext1: $F_{9,144}$=7.063, p<0.0001, Ext2: $F_{9,144}$=5.472, p<0.0001) with no significant group difference (Ext1: $F_{1,16}$=1.163, p=0.2968, Ext2: $F_{1,16}$=0.8747, p=0.3636; **Figure 3h**). In extinction retrieval tests, photostimulation via pulses of blue light (10 Hz, 20 ms duration), which evoked action potential firing with high temporal precision and spike fidelity (97.7 ± 2.3%; **Supplementary Figure 4**), was delivered to the IL during the first four trial blocks of the CS presentation. Tone-induced freezing in ChR2-expressing mice was robustly reduced compared to that in control mice during the light-ON period **(**significant group × trial interaction, $F_{1,16}$=14.2, p<0.01, two-way repeated-measures ANOVA; p<0.01, Bonferroni *post hoc* test; **Figure 3i**). Notably, this reduced freezing in the ChR2 group was similar to baseline pre-CS freezing measured before tone presentation, indicating almost complete block of conditioned freezing (significant group × trial interaction, $F_{1,16}$=9.599, p<0.01, two-way repeated-measures ANOVA; p=0.1569, Bonferroni *post hoc* test; **Figure 3i**). The enhanced expression of fear extinction induced by IL photoactivation was reversible and specific, as indicated by return of freezing behavior to control levels once the light was turned off (39.5% and 41.2% for EGFP and ChR2, respectively; **Figure 3i**). Taken together, these results reveal that increasing the activity of IL excitatory neurons at the time of extinction retrieval is sufficient to enhance expression of fear extinction. These results, in conjunction with our observations of opposite effects produced by photoinhibiting IL, indicate that IL activity plays a causal role in controlling fear expression after extinction.

Photostimulation of IL excitatory neurons does not affect expression of conditioned fear before extinction

If CS-induced activation of IL is a critical mechanism for fear inhibition after extinction, artificially activating IL neurons alone may be sufficient to induce extinction-like inhibition

of conditioned fear expression even without extinction training. In fact, previous studies reported that electrical stimulation of the medial prefrontal cortex, or specifically the IL, reduces conditioned freezing in animals that do not undergo extinction (Milad *et al*, 2002; Milad *et al*, 2004; Vidal-Gonzalez *et al*, 2006). However, because uncertainties remain because of potential non-specific effects of electrical stimulation, we took advantage of ChR2-mediated photostimulation to test this prediction. During the fear memory test, ChR2-expressing neurons in the IL were photostimulated using the same conditions as before and were paired with tone presentation (**Figure 4a**). In contrast to previous findings, we found that conditioned freezing in ChR2 mice was not significantly altered by IL photoactivation compared with control mice (**Figure 4b**; p=0.50, Student's t test).

It is possible that the light-stimulation conditions used (10 Hz, 20 ms pulse duration) were too weak or physiologically irrelevant to induce behavioral changes under no-extinction conditions. Indeed, it has been reported that IL activity is increased up to a maximal rate of 20 Hz in the first 100–400 ms after tone onset during extinction recall (Milad *et al*, 2002), and electrical stimulation is effective in reducing conditioned freezing in a temporally specific manner (Milad *et al*, 2004). Based on these findings, we next tested photostimulation at 20 Hz (10 ms pulse duration), with or without a 0.1 s delay after tone onset. Patch-clamp recordings of brain slices confirmed that reliable spike firing (85.6 ± 6.1%) was evoked with such stimuli (**Supplementary figure 3**). Again, we found no significant group difference in freezing without a delay (**Figure 4d**; p=0.2815, Student's t test) or with it (**Figure 4e**; p=0.1385, Student's t test), although ChR2 mice showed slightly reduced freezing, compared to control mice, upon IL photoactivation (**Figures 4d, e**). Finally, we tested bilateral stimulation of the IL. We observed no significant difference in freezing between control and ChR2 groups (p=0.60, Student's t test; **Figure 4f**), indicating no effect of bilateral IL

activation on conditioned fear expression. Therefore, these results show that activation of IL excitatory neurons before extinction does not affect expression of conditioned fear, suggesting that IL activation alone is not sufficient to suppress the expression of conditioned fear.

Given the lack of effect of IL stimulation on non-extinguished conditioned fear, we next asked whether the effect of IL stimulation after extinction of auditory fear memory is specific to that fear memory. To test this idea, we investigated the effect of IL photostimulation on the expression of contextual fear that did not undergo extinction after extinction training of auditory fear. After auditory fear conditioning and extinction training to the tone CS as before, mice were re-introduced to the conditioned context where they received shock during fear conditioning (**Figure 5a, b**). There were no significant differences in freezing between control and ChR2 groups during testing in the light-ON or light-OFF periods (group × trial interaction, $F_{1,13}$=0.96, p=0.35; **Figure 5c**). Therefore, our results show that IL stimulation effect is specific to the fear memory that undergoes extinction.

Discussion

We used optogenetics to determine that precise, real-time inhibition of activity in IL, but not in PL, impairs the expression of fear extinction. Further, activation of excitatory neurons in the IL enhances expression of fear extinction, while neither manipulation of IL activity significantly affects the expression of conditioned fear. Therefore, we conclude that IL activity acts as an important top-down control for fear expression after extinction that is necessary and sufficient to inhibit expression of conditioned fear after extinction.

Because there are no known genetic targeting approaches specific to the IL, we relied

on spatially precise injections of virus for the delivery of opsins (eNpHR3.0 or ChR2) specifically to the IL. For highly restricted expression of eNpHR3.0 or ChR2 in the IL with minimal damage to other vmPFC areas, we optimized virus delivery conditions, including target coordinates, virus injection volume, and virus titer. Virus expression was predominantly restricted to neurons in the IL in most animals included in our data analysis, although infections were occasionally observed in the ventral part of the PL and DP. Thus, our optogenetic manipulations were highly specific to cells in the IL, although possible contributions of other vmPFC regions, such as the PL and DP, cannot be completely ruled out.

Our current findings are in line with the 'IL hypothesis', which posits that the IL plays a critical role in the control of fear expression after extinction (Herry *et al*, 2010; Quirk *et al*, 2006; Quirk and Mueller, 2008; Sotres-Bayon *et al*, 2004). Among the previous evidence in support of the IL hypothesis, perhaps the best is the demonstration that the CS-evoked response of IL neurons *in vivo* is increased during recall of extinction memory (Milad *et al*, 2002). This observation suggests that IL activity controls fear expression after extinction such that its activation by the CS mediates the expression of fear extinction. However, this causal link has remained elusive, mainly because techniques capable of providing the necessary spatiotemporal resolution for control of IL activity have been lacking. In addition, most previous studies have focused largely on extinction acquisition or consolidation processes and not specifically on the expression of fear extinction. Accordingly, in most cases functional manipulations such as lesions, drug treatment, or electrical stimulation were performed before fear conditioning or during extinction training (Milad *et al*, 2002; Milad *et al*, 2004; Morgan *et al*, 1995; Morgan *et al*, 1993; Quirk *et al*, 2000). In a few studies, lesions or pharmacological interventions were applied after extinction training to examine effects on extinction retrieval, but such manipulations were targeted to the mPFC

generally rather than specifically to the IL. Moreover, some of these studies failed to support a role for the IL in extinction recall. For example, one study showed that lesions made in the vmPFC 1 day after extinction training had no effect on normal expression of extinction memory (Garcia *et al*, 2006). In another study, inactivation of the vmPFC by the sodium channel blocker tetrodotoxin, infused 30 min prior to the extinction recall test, actually reduced freezing (Sierra-Mercado *et al*, 2006). These findings are inconsistent with our data. One possible explanation for these discrepancies is that lesion- or drug-mediated inactivation of the vmPFC encompassing both the IL and PL produces compound effects. Differences in experimental conditions, such as temporal relationships, may also be important contributors. For example, the lesion study employed a 7-day interval between extinction training and testing (Garcia *et al*, 2006), whereas we tested animals 1 day after extinction training. Interpretation of data may also be further complicated by undesired effects produced by the relatively poor spatial and temporal precision of drug application and potential compensation by other brain structures in lesion or drug-infusion protocols, as shown in a recent study (Goshen *et al*, 2011). A recent study reported that optogenetic silencing of IL activity has no effect on retrieval of extinction of auditory conditioned fear (Do-Monte *et al*, 2015), which is inconsistent with our conclusion. ~~While we have no clear explanation for the discrepancy between the two sets of results, there are several differences in experimental conditions between the two studies that might account for the discrepancy. While Do-Monte *et al*. (2015) employed rats and bilaterally silenced IL activity, we used mice and silenced IL unilaterally. In our case, the guide cannula was angled toward the midline to avoid potential damage to PL, while it was vertically implanted in the rat study. Another~~ One of major differences between two studies is that we used the hSyn promoter to express eNpHR3.0 that targets both glutamatergic and GABAergic neurons, while Do-Monte *et al*. (2015) used the CaMKIIα

promoter to selectively silence glutamatergic neurons. Because we also found no significant effect of IL photoinhibition when NpHR expression was driven by CaMKIIα promoter under our conditions, we presume that manipulating different populations of neurons via different promoters was the main reason for the different results of the two studies, implying that GABAergic neurons in the IL might also critically participate in regulating extinction retrieval. Notwithstanding, our positive data support the idea that IL activity is essential for expression of fear extinction at the time of retrieval.

Given the previous reports showing that IL activation – via electrical, pharmacological or optogenetic stimulation - reduces conditioned freezing, even in animals that did not undergo extinction training (Do-Monte *et al*, 2015; Milad *et al*, 2002; Milad *et al*, 2004; Thompson *et al*, 2010; Vidal-Gonzalez *et al*, 2006), we anticipated that IL photostimulation might produce similar effects. However, we unexpectedly found that IL photoactivation did not affect the expression of conditioned freezing without extinction (**Figure 4**). We tested different light stimulation frequencies, with or without a delay after tone onset, as well as bilateral stimulation. But none of the conditions that we tested produced a significant reduction in conditioned freezing, even though they did have other behavioral effects. Unlike electrical or pharmacological stimulation, the optogenetic photostimulation paradigm used here was specific to CaMKIIα-positive neurons in the IL. Thus, it is possible that stimulation of a larger population of cells encompassing different types of cells within the IL may be necessary to induce inhibition of conditioned freezing. However, this is less likely, given a recent optogenetic study in rat showing that the same optical activation of CaMKIIα-positive neurons induced a strong reduction in conditioned freezing (Do-Monte *et al*, 2015). Alternatively, it is possible that photostimulation was weaker in our conditions, so that an insufficient number of neurons were activated to produce a behavioral outcome.

However, this possibility seems unlikely because the same photostimuli robustly reduced conditioned freezing after extinction. Thus, although it is unclear how to explain the discrepancy, our data strongly indicate that an increase in IL activity alone is insufficient, or at least is less efficient, for inhibiting fear output in the absence of extinction. Importantly, the expression of contextual fear was not affected by IL photostimulation after extinction of auditory fear. This result further supports that control of fear expression by IL is specific to the fear memory that undergoes extinction. Taken together, our data support the hypothesis that CS-evoked activity in the IL is a key circuit mechanism in the top-down control of fear expression after extinction, but imply that plastic changes in broader neural networks involved in fear and fear extinction are also required for the expression of fear extinction (Amano *et al*, 2010; Cho *et al*, 2013; Rosenkranz *et al*, 2003; Senn *et al*, 2014; Sierra-Mercado *et al*, 2011; Trouche *et al*, 2013).

IL activity correlates with a decrement in freezing behavior during retrieval of extinction (Milad *et al*, 2002; Milad *et al*, 2004; Vidal-Gonzalez *et al*, 2006). Consistent with this, we observed that photostimulation-induced enhancement of IL activity after extinction elicited a further reduction in fear expression to as low as to pre-CS levels (**Figure 3i**). This suggests that extinction training itself may not induce full activation of the IL. Thus, modulation of the intrinsic excitability of the IL excitatory neurons or activity of local inhibitory networks within IL may provide another regulation point for the fear extinction (Herry *et al*, 1999; Senn *et al*, 2014). How the intrinsic plasticity of IL neurons is regulated by the internal state of the brain or by experience will be an important issue to consider in future studies.

In summary, by bidirectionally manipulating IL activity with precise spatiotemporal resolution using optogenetics during extinction memory retrieval, our study overcomes the

limitations of previous research and provides compelling evidence establishing a causal role for IL activity in controlling the expression of fear extinction. The level of IL activity determines fear output expressed to the CS after extinction, providing important insights into how vmPFC controls fear expression as a top-down mechanism in extinction.

Funding and Disclosure

This work was supported by grants from the National Research Foundation of Korea (NRF) funded by the Korea government (MSIP) (2011-0013173 and 2014R1A2A1A10053821), the KAIST Future Systems Healthcare Project, the Brain Research Program through the National Research Foundation of Korea(NRF) (2015M3C7A1027351), in part by a grant from the Intelligent Synthetic Biology Center of the Global Frontier Project funded by the Ministry of Science, ICT & Future Planning (2011-0031955), by a grant from the World Class Institute (WCI) program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (MEST) (No. WCI 2009-003) and by a CRP grant from the National Research Foundation of Singapore. The authors declare no conflict of interest.

Acknowledgements

We thank members of our laboratories for critical discussions and helpful comments. Special thanks go to Dr. R. Nakajima for helpful advice on electrophysiological techniques. J.-H.H. designed and directed the study. H.-S.K. designed and generated AAV viruses, and performed stereotaxic brain surgery, in vitro brain slice patch-clamp recordings, immunohistochemistry, confocal imaging, and animal behavior experiments. H.-Y.C. performed manual scoring to quantify freezing behavior and histological verification for

behavior experiments. G.J.A. contributed to the design of in vitro brain slice patch-clamp experiments. H.-S.K., G.J.A. and J.-H.H. analyzed data and wrote the manuscript.

Supplementary information is available at the *Neuropsychopharmacology* website

References

Amano T, Unal CT, Pare D (2010). Synaptic correlates of fear extinction in the amygdala. *Nature neuroscience* **13**: 489–494.

Benson DL, Isackson PJ, Gall CM, Jones EG (1992). Contrasting patterns in the localization of glutamic acid decarboxylase and Ca^{2+}/calmodulin protein kinase gene expression in the rat central nervous system. *Neuroscience* **46**: 825–849.

Chang CH, Berke JD, Maren S (2010). Single-unit activity in the medial prefrontal cortex during immediate and delayed extinction of fear in rats. *PLoS ONE* **5**: e11971.

Cho JH, Deisseroth K, Bolshakov VY (2013). Synaptic encoding of fear extinction in mPFC-amygdala circuits. *Neuron* **80**: 1491–1507.

Do-Monte FH, Manzano-Nieves G, Quinones-Laracuente K, Ramos-Medina L, Quirk GJ (2015). Revisiting the role of infralimbic cortex in fear extinction with optogenetics. *J Neurosci* **35**: 3607–3615.

Fanselow MS, Gale GD (2003). The amygdala, fear, and memory. *Ann NY Acad Sci* **985**: 125–134

Farinelli M, Deschaux O, Hugues S, Thevenet A, Garcia R (2006). Hippocampal train stimulation modulates recall of fear extinction independently of prefrontal cortex synaptic plasticity and lesions. *Learn Mem* **13**(3): 329–334.

Franklin KBJ, Paxinos G (1997). *The mouse brain in stereotaxic coordinates*. Academic Press: San Diego, CA.

Garcia R, Chang CH, Maren S (2006). Electrolytic lesions of the medial prefrontal cortex do not interfere with long-term memory of extinction of conditioned fear. *Learn Mem* **13**: 14–17.

Gewirtz JC, Falls WA, Davis M (1997). Normal conditioning inhibition and extinction of freezing and fear-potentiated startle following electrolytic lesions of medial prefrontal cortex in rats. *Behav Neurosci* **111**: 712–726.

Goshen I, Brodsky M, Prakash R, Wallace J, Gradinaru V, Ramakrishnan C *et al* (2011). Dynamics of retrieval strategies for remote memories. *Cell* **147**: 678–689.

Herry C, Ferraguti F, Singewald N, Letzkus JJ, Ehrlich I, Luthi A (2010). Neuronal circuits of fear extinction. *Eur J Neurosci* **31**: 599–612.

Herry C, Garcia R (2002). Prefrontal cortex long-term potentiation, but not long-term depression, is associated with the maintenance of extinction of learned fear in mice. *J Neurosci* **22**: 577–583.

Herry C, Vouimba RM, Garcia R (1999). Plasticity in the mediodorsal thalamo-prefrontal cortical transmission in behaving mice. *J Neurophysiol* **82**: 2827–2832.

Hugues S, Chessel A, Lena I, Marsault R, Garcia R (2006). Prefrontal infusion of PD098059 immediately after fear extinction training blocks extinction-associated prefrontal synaptic plasticity and decreases prefrontal ERK2 phosphorylation. *Synapse* **60**: 280–287.

Kapp BS, Frysinger RC, Gallagher M, Haselton JR (1979). Amygdala central nucleus lesions: effect on heart rate conditioning in the rabbit. *Physiol Behav* **23**: 1109–1117.

Kwon JT, Nakajima R, Kim HS, Jeong Y, Augustine GJ, Han JH (2014). Optogenetic activation of presynaptic inputs in lateral amygdala forms associative fear memory. *Learn Mem* **21**: 627–633.

LaLumiere RT, Niehoff KE, Kalivas PW (2010). The infralimbic cortex regulates the consolidation of extinction after cocaine self-administration. *Learn Mem* **17**: 168–175.

Laurent V, Westbrook RF (2009). Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. *Learn Mem* **16**: 520–529.

Lebron K, Milad MR, Quirk GJ (2004). Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. *Learn Mem* **11**: 544–548.

LeDoux JE (2000). Emotion circuits in the brain. *Ann Rev Neurosci* **23**: 155–184.

Likhtik E, Popa D, Apergis-Schoute J, Fidacaro GA, Pare D (2008). Amygdala intercalated neurons are required for expression of fear extinction. *Nature* **454**: 642–645.

Liu XB, Jones EG (1996). Localization of alpha type II calcium calmodulin-dependent protein kinase at glutamatergic but not gamma-aminobutyric acid (GABAergic) synapses in thalamus and cerebral cortex. *Proc Natl Acad Sci USA* **93**: 7332–7336.

Milad MR, Quirk GJ (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**: 70–74.

Milad MR, Quirk GJ (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Ann Rev Psycho* **63**: 129–151.

Milad MR, Rauch SL, Pitman RK, Quirk GJ (2006). Fear extinction in rats: implications for human brain imaging and anxiety disorders. *Biol Psycho* **73**: 61–71.

Milad MR, Vidal-Gonzalez I, Quirk GJ (2004). Electrical stimulation of medial prefrontal cortex reduces conditioned fear in a temporally specific manner. *Behav Neurosci* **118**: 389–394.

Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol Psychiatry* **62**: 446–454.

Morgan MA, LeDoux JE (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav Neurosci* **109**: 681–688.

Morgan MA, Romanski LM, LeDoux JE (1993). Extinction of emotional learning: contribution of medial prefrontal cortex. *Neurosci lett* **163**: 109–113.

Morgan MA, Schulkin J, LeDoux JE (2003). Ventral medial prefrontal cortex and emotional perseveration: the memory for prior extinction training. *Behav Brain Res* **146**: 121–130.

Myers KM, Davis M (2007). Mechanisms of fear extinction. *Mol psychiatry* **12**: 120-150.

Pavlov I (1927). *Conditioned Reflexes.* Oxford University Press: Oxford, UK.

Peters J, Kalivas PW, Quirk GJ (2009). Extinction circuits for fear and addiction overlap in prefrontal cortex. *Learn Mem* **16**: 279–288.

Phelps EA, Delgado MR, Nearing KI, LeDoux JE (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron* **43**: 897–905.

Quirk GJ, Garcia R, Gonzalez-Lima F (2006). Prefrontal mechanisms in extinction of conditioned fear. *Biol Psychiatry* **60**: 337–343.

Quirk GJ, Mueller D (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology* **33**: 56–72.

Quirk GJ, Russo GK, Barron JL, Lebron K (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *J Neurosci* **20**: 6225–6231.

Rosenkranz JA, Moore H, Grace AA (2003). The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *J Neurosci* **23**: 11054–11064.

Santini E, Quirk GJ, Porter JT (2008). Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. *J Neurosci* **28**: 4028–4036.

Senn V, Wolff SB, Herry C, Grenier F, Ehrlich I, Grundemann J *et al* (2014). Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* **81**: 428–437.

Sierra-Mercado D, Corcoran KA, Lebron-Milad K, Quirk GJ (2006). Inactivation of the ventromedial prefrontal cortex reduces expression of conditioned fear and impairs subsequent recall of extinction. *Eur J Neurosci* **24**: 1751–1758.

Sierra-Mercado D, Padilla-Coreano N, Quirk GJ (2011). Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology* **36**: 529–538.

Sotres-Bayon F, Bush DE, LeDoux JE (2004). Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn Mem* **11**: 525–535.

Sotres-Bayon F, Quirk GJ (2010). Prefrontal control of fear: more than just extinction. *Curr Opin Neurobiol* **20**: 231–235.

Thompson BM, Baratta MV, Biedenkapp JC, Rudy JW, Watkins LR, Maier SF (2010). Activation of the infralimbic cortex in a fear context enhances extinction learning. *Learn Mem* **17**: 591–599.

Trouche S, Sasaki JM, Tu T, Reijmers LG (2013). Fear extinction causes target-specific remodeling of perisomatic inhibitory synapses. *Neuron* **80**: 1054–1065.

Van den Oever MC, Rotaru DC, Heinsbroek JA, Gouwenberg Y, Deisseroth K, Stuber GD *et al* (2013). Ventromedial prefrontal cortex pyramidal cells have a temporal dynamic role in recall and extinction of cocaine-associated memory. *J Neurosci* **33**: 18225–18233.

Vidal-Gonzalez I, Vidal-Gonzalez B, Rauch SL, Quirk GJ (2006). Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. *Learn Mem* **13**: 728–733.

Willcocks AL, McNally GP (2013). The role of medial prefrontal cortex in extinction and

reinstatement of alcohol-seeking in rats. *Eur J Neurosci* **37**: 259–268.

Figure legends

Figure 1 Photoinactivation of IL impaired expression of fear extinction with no effect on the expression of conditioned fear before extinction. (**a**) Left: AAV viral vector construct for expressing eNpHR3.0-EYFP under the control of the human synapsin promoter, and schematic depiction of the experimental design for selective optogenetic inhibition of IL neuronal activity by eNpHR3.0. The vertical line indicates the anterior–posterior axis of the virus injection site; the sagittal and coronal section is shown below. The cannula was implanted 15 degrees to the dorsoventral axis to avoid damaging the dorsal part of the vmPFC. Right: Representative confocal microscopic image showing specific expression of eNpHR3.0 protein in the IL, visualized by EYFP fluorescence. Insets, magnified images showing membrane-targeted eNpHR3.0 protein expression. (**b**) Histological confirmation of virus expression and optic fiber placement (° represents the location of the fiber optic tip).

Representative viral infection areas in the IL for the largest (dark gray) and smallest (light gray) area of eNpHR3.0-EYFP expression. (**c**) Sample photocurrent trace from a brain slice patch-clamp recording of EYFP-positive neurons in the IL upon green light illumination. Light of different intensities induced graded changes in photocurrents. (**d**) Relationship between photocurrent amplitude and light intensity (2-s light exposure; $n = 14$). (**e**) Sample trace of spike firing recorded in EYFP-positive neurons in the IL. Increasing light intensity led to stronger inhibition of spike firing evoked by 200-pA current injection. (**f**) Graph showing the relationship between percentage ratio of spike firing inhibition and different light intensities (2-s light exposure; $n = 12$). (**g**) Experimental procedure. Mice injected with either eNpHR3.0 or EGFP control AAV viral vector into the IL were trained for auditory fear conditioning and subsequently administered extinction training. Green light (561 nm) was delivered to inactivate IL activity. (**h**) The percentage of freezing during auditory fear conditioning (FC) and extinction training (Ext). (**i**) The percentage of freezing during extinction retrieval. Light illumination was applied during the first half of total tone CS presentations and was turned off during the remaining half. Mice from the eNpHR3.0 group showed greater freezing behavior compared with the control group during the light-ON period and this difference disappeared when lights were turned off (light-OFF) ($n = 9$ for each group). $*p<0.05$ (*post-hoc* comparison). (**j**) The percentage of freezing exhibited in the absence of tone with (light-ON) or without (light-OFF) light illumination of the IL. There was no significant group difference in freezing ($n = 10$ for each group). (**k**) Experimental procedure. Mice injected with eNpHR3.0 into the IL were trained for auditory fear conditioning with a single pairing of tone and shock, and tested 24 h later. Freezing was determined in the presence of tone for 2 min. Light with a wavelength of 561 nm was delivered during the first 1 min of tone presentation and then turned off for the next 1 min.

The percentage of freezing during fear memory test with (light-ON) or without (light-OFF) photoinhibition of the IL. Mice from the two groups showed a similar level of freezing in response to tone presentation regardless of light illumination (EGFP, $n = 5$; eNpHR3.0, $n = 7$). Data in **d, e, g–k** are expressed as means ± s.e.m. x, baseline (pre-CS) freezing levels.

Figure 2 Photoinactivation of PL did not affect expression of fear extinction. (**a**) Top: Schematic depiction of the experimental design for AAV virus-mediated eNpHR3.0 expression in the PL. Bottom: Confocal microscopic image showing expression of eNpHR3.0 protein fused with EYFP in the PL (green). Insets, magnified images of eNpHR3.0-EYFP–expressing neurons in the PL. (**b**) Experimental procedure. Mice injected with either eNpHR3.0 or control EGFP AAV viral vector into the PL were trained for auditory fear conditioning and subsequently administered extinction training. Green light (561 nm) was delivered as before to inactivate the PL during the first half of total tone presentations. (**c**) The percentage of freezing during auditory fear conditioning (FC) and extinction training procedure (Ext). (**d**) The percentage of freezing during extinction retrieval. Mice from eNpHR3.0 and control EGFP groups displayed no significant differences in freezing throughout tone presentations, even with light illumination ($n = 9$ for each group). Data in **c**, **d** are expressed as means ± s.e.m. x, baseline (pre-CS) freezing levels.

Figure 3 Photoactivation of IL excitatory neurons enhanced expression of fear extinction. (**a**) Left: AAV vector construct for expressing ChR2-Venus in excitatory neurons in the IL, and schematic depiction of the experimental design for photoactivation of the IL. Mice were injected with AAV-CaMKIIα-ChR2-Venus and implanted with a guide cannula targeting the IL at a 15-degree angle to the dorsoventral axis. Right: Representative confocal microscopic

image showing the expression pattern of Venus-tagged ChR2 (green). Insets, magnified images of ChR2-expressing cells in the IL. (**b**) Representative confocal microscopic images showing selective expression of ChR2-Venus in CaMKIIα-positive neurons in the IL. (**c**) Sample photocurrent trace from a brain slice patch-clamp recording of Venus-positive neurons in the IL upon blue light illumination. Light of different intensities induced graded changes in photocurrents. (**d**) Bar graph showing photocurrent amplitudes (steady state) evoked by different light intensities (2-s duration; n = 13). (**e**) Sample traces (black line) from brain slice patch-clamp recordings of ChR2-Venus–expressing excitatory neurons in the IL. Spike firing responses to different light frequencies (blue traces; 4.6 mW mm^{-2}, 4-ms duration) are shown. (**f**) Mean probability of evoked action potentials by different light frequencies. Probability decreased with increasing light pulse frequency (n = 12). (**g**) Experimental procedure. Mice injected with either ChR2-Venus or EGFP vector into the IL were trained for fear conditioning and subsequently administered extinction training. During extinction memory tests, tone-induced freezing was determined with and without blue light (473 nm) illumination at a frequency of 10 Hz. (**h**) The percentage of freezing during auditory fear conditioning (FC) and extinction training procedure (Ext). (**i**) The percentage of freezing during extinction memory tests. Mice injected with ChR2-Venus exhibited significantly reduced freezing to the tone CS compared with EGFP control mice following photoactivation of IL excitatory neurons (light-ON; n = 9 for each group). $^{**}P < 0.01$ (*post-hoc* comparison). Data in **d, f, h, i** are expressed as means ± s.e.m. x, baseline (pre-CS) freezing levels.

Figure 4 Photoactivation of IL did not affect the expression of conditioned fear without extinction. (**a**) Experimental procedure. Mice injected with either ChR2-Venus or control

EGFP into the IL were trained for auditory fear conditioning with single pairing of tone and shock, and then tested 24 h later. During tone presentation, a 10-Hz light stimulation was delivered to the IL simultaneously with tone onset. (**b**) The percent of freezing during fear memory tests. There was no significant group difference in freezing between EGFP (n = 8) and ChR2 (n = 9) groups. (**c**) Experimental procedure. Mice injected with either ChR2-Venus or control EGFP into the IL were trained for auditory fear conditioning with three pairings of tone and shock, and then tested 24 h later. Three different photoactivation conditions were used: 20 Hz with no delay after tone onset (EGFP, n = 7; ChR2, n = 8) (**d**), 20 Hz with a 0.1-s delay (EGFP, n = 10; ChR2, n = 7) (**e**), and a 10-Hz bilateral IL stimulation with no delay (EGFP, n = 11; ChR2, n = 8) (**f**). (**d, e, f**) The percentage of freezing during tone CS presentation is shown. There was no significant difference in freezing between ChR2-Venus and EGFP control mice. Data in **b**, **d**–**f** are expressed as means ± s.e.m.

Figure 5 Photoactivation of IL did not affect the expression of contextual fear after extinction of auditory conditioned fear. (**a**) Experimental procedure. Mice injected with either ChR2 or control EGFP AAV viral vector into the IL were trained for auditory fear conditioning and subsequently administered extinction training. Blue light (473 nm) was delivered as before to activate the IL during the first minute of 2-min contextual fear memory test. (**b**) The percentage of freezing during auditory fear conditioning (FC) and extinction training procedure (Ext). (**c**) The percentage of freezing during contextual fear memory tests. Mice injected with either ChR2-Venus (n = 6) or control EGFP (n = 9) into the IL did not show significant difference in freezing. Data in **b**, **c** are expressed as means ± s.e.m. x, baseline (pre-CS) freezing levels.