

# Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression

Julien Courtin<sup>1,2</sup>, Fabrice Chaudun<sup>1,2</sup>, Robert R. Rozeske<sup>1,2</sup>, Nikolaos Karalis<sup>1,2</sup>, Cecilia Gonzalez-Campo<sup>1,2</sup>, H     Wurtz<sup>1,2</sup>, Azzedine Abdi<sup>3,4</sup>, Jerome Baufreton<sup>3,4</sup>, Thomas C. M. Bienvenu<sup>1,2</sup> & Cyril Herry<sup>1,2</sup>

**Synchronization of spiking activity in neuronal networks is a fundamental process that enables the precise transmission of information to drive behavioural responses<sup>1–3</sup>. In cortical areas, synchronization of principal-neuron spiking activity is an effective mechanism for information coding that is regulated by GABA ( $\gamma$ -aminobutyric acid)-ergic interneurons through the generation of neuronal oscillations<sup>4,5</sup>. Although neuronal synchrony has been demonstrated to be crucial for sensory, motor and cognitive processing<sup>6–8</sup>, it has not been investigated at the level of defined circuits involved in the control of emotional behaviour. Converging evidence indicates that fear behaviour is regulated by the dorsomedial prefrontal cortex<sup>9–12</sup> (dmPFC). This control over fear behaviour relies on the activation of specific prefrontal projections to the basolateral complex of the amygdala (BLA), a structure that encodes associative fear memories<sup>13–15</sup>. However, it remains to be established how the precise temporal control of fear behaviour is achieved at the level of prefrontal circuits. Here we use single-unit recordings and optogenetic manipulations in behaving mice to show that fear expression is causally related to the phasic inhibition of prefrontal parvalbumin interneurons (PVINs). Inhibition of PVIN activity disinhibits prefrontal projection neurons and synchronizes their firing by resetting local theta oscillations, leading to fear expression. Our results identify two complementary neuronal mechanisms mediated by PVINs that precisely coordinate and enhance the neuronal activity of prefrontal projection neurons to drive fear expression.**

To identify the prefrontal circuitry involved in conditioned fear behaviour, mice were implanted with recording electrodes aimed at the dmPFC, and submitted to auditory fear conditioning, a robust learning paradigm in which animals learn to associate a neutral stimulus (the conditioned stimulus, CS) with a coincident aversive foot-shock (the unconditioned stimulus, US) (Fig. 1a). Re-exposure to the CS induces the expression of various conditioned fear responses, including an immobilization reaction called freezing. Twenty-four hours after conditioning, mice displayed a selective increase in freezing during presentations of the CS associated with the US (CS<sup>+</sup>), which returned to baseline levels by the end of the second extinction session (Fig. 1b). One week later, CS<sup>+</sup> presentations induced a selective fear recovery (Fig. 1b). Among the 732 neurons recorded in dmPFC, 493 (67.3%) displayed significant excitatory or inhibitory phasic responses to CS<sup>+</sup> presentations following conditioning. To dissect dmPFC circuits involved in the control of fear behaviour, we separated the CS<sup>+</sup>-responsive neurons into putative principal neurons (PNs,  $n = 351$ ) and interneurons (INs,  $n = 142$ ) using unsupervised clustering and cross-correlogram analyses (Extended Data Fig. 1). Among dmPFC INs, principal component analyses revealed two main subclasses with opposite CS-evoked responses during fear expression (Fig. 1c, d, and Extended Data Fig. 2a, b). Type 1 INs ( $n = 68$ ) displayed short-latency, CS-evoked activity correlated with high (CS<sup>+</sup>), but not low (CS<sup>−</sup>), fear states. Conversely, type 2 INs ( $n = 15$ ) were strongly inhibited during high but not low fear states (Fig. 1c, d). Correlation analyses carried out

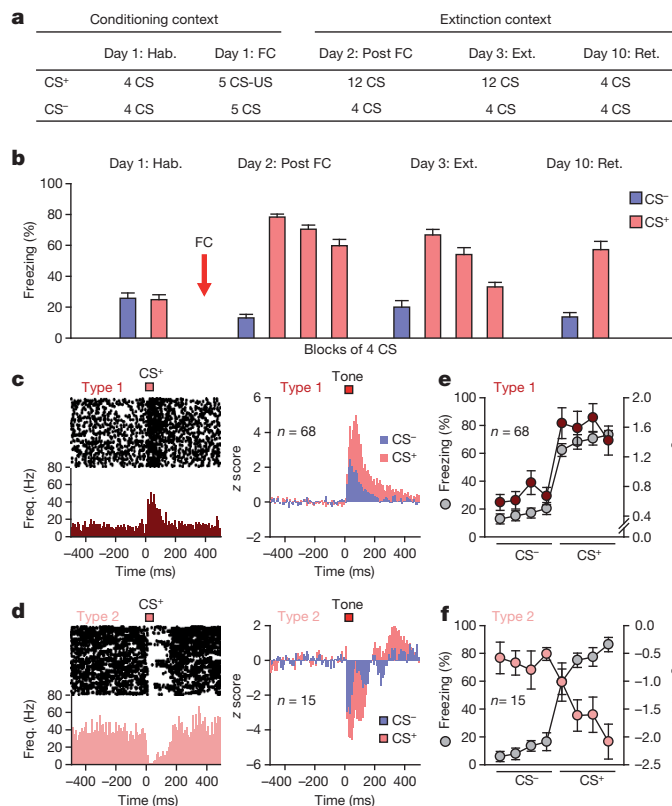
between changes in activity after CS presentations and freezing levels revealed that the firing of type 1 and type 2 INs were correlated and inversely correlated, respectively, with freezing (Fig. 1e, f). Moreover, latency and cross-correlation analyses of simultaneously recorded cells revealed that CS<sup>+</sup>-evoked excitation of type 1 INs preceded type 2 INs CS<sup>+</sup>-evoked inhibition (Extended Data Fig. 2c–e).

Interestingly, whereas type 1 INs displayed moderate firing rates ( $16.2 \pm 1.5$  Hz) and were weakly modulated with local theta oscillations, type 2 INs showed fast firing activity ( $43.9 \pm 9.7$  Hz) and were strongly modulated with local theta, suggesting that type 2 INs are PVINs<sup>16</sup> (Extended Data Fig. 2f–h). To address this possibility, we selectively infected PVINs with injections of a conditional adeno-associated virus (AAV) encoding for archaeorhodopsin in the dmPFC of mice expressing the Cre recombinase under the control of a PV promoter (PV-IRES-Cre; Fig. 2a and Extended Data Fig. 3a, b). Using this strategy, we optically silenced the firing of type 2 ( $n = 5/5$  (5 out of 5)) but not type 1 INs ( $n = 0/9$ ), indicating that type 2 INs belong to the PVIN population (Fig. 2b). Remarkably, among light-reactive PVINs ( $n = 9$ ), only type 2 PVINs ( $n = 5$ ) displayed significant decreases in CS-evoked activity following conditioning, suggesting a functional role of this subpopulation during fear behaviour (Extended Data Fig. 4a–d). In summary, we identified two subclasses of dmPFC INs whose activities oppositely correlate with fear behaviour and demonstrated that type 2 INs are PVINs.

To determine whether the CS-evoked inhibition of type 2 PVINs causes fear expression, PV-IRES-Cre mice received intra-dmPFC injections of a conditional AAV encoding for archaeorhodopsin or channelrhodopsin. Infection of dmPFC PVINs did not change their electrophysiological characteristics (Extended Data Fig. 3c–e). Before fear conditioning, optical silencing of PVINs induced freezing (Fig. 2c). Moreover, after fear extinction, CS<sup>+</sup> presentations coupled to optical silencing of PVINs, including type 2 INs, consistently reinstated fear responses (Fig. 2c and Extended Data Fig. 4c–e). Conversely, optical activation of PVINs transiently inhibited freezing (Fig. 2d). To control that freezing induced by CS-evoked inhibition of type 2 INs did not result from motor impairments, we optically inhibited PVINs during a place avoidance paradigm, in which mice could actively avoid the compartment in which they received optical silencing. Under these conditions, optogenetic silencing of PVINs produced place aversion relative to control animals (Extended Data Fig. 5). These data demonstrate that fear expression is causally related to the inhibition of dmPFC PVINs, including type 2 INs.

PVINs target the perisomatic region of PNs, thereby providing powerful inhibition of dmPFC output activity<sup>17</sup>. Therefore, CS<sup>+</sup>-evoked inhibition of PVINs during fear behaviour might disinhibit PNs, a permissive mechanism that would gate fear responses. Consistent with this, the vast majority of tone-reactive PNs ( $n = 308/351$ , 87.7%) significantly increased their activity upon CS<sup>+</sup> relative to CS<sup>−</sup> presentations (Fig. 3a). Moreover, the optogenetic activation of PVINs inhibited PNs, prevented CS<sup>+</sup>-induced activation of PNs and reduced freezing (Extended Data Fig. 6a–c). Conversely, light-induced inhibition of

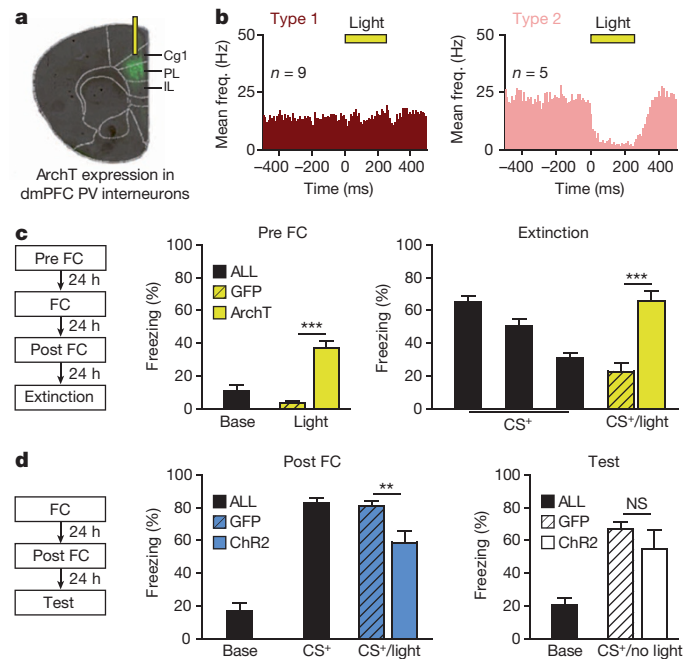
<sup>1</sup>INSERM, Neurocentre Magendie, U862, 146 Rue L  o-Saignat, Bordeaux 33077, France. <sup>2</sup>University of Bordeaux, Neurocentre Magendie, U862, 146 Rue L  o-Saignat, Bordeaux 33077, France. <sup>3</sup>University of Bordeaux, Institut des Maladies Neurod  g  n  ratives, UMR 5293, Bordeaux F-33000, France. <sup>4</sup>CNRS, Institut des Maladies Neurod  g  n  ratives, UMR 5293, Bordeaux F-33000, France.



**Figure 1 | Firing of distinct dmPFC INs oppositely correlates with fear expression.** **a**, Protocol. **b**, During habituation (Hab.), mice ( $n = 29$ ) exhibited low freezing during CS<sup>-</sup> and CS<sup>+</sup>. After fear conditioning (Post FC; the first extinction session), CS<sup>+</sup> (CS presentations 1–12, grouped into blocks of 4) induced high freezing (Wilcoxon signed-rank tests, CS<sup>-</sup> versus each CS<sup>+</sup> block; all  $P < 0.001$ ). After extinction (Ext. (the second extinction session),  $n = 28$  mice), CS<sup>+</sup> (CS<sup>+</sup> 9–12; the last 4 CS<sup>+</sup> of the extinction session) and CS<sup>-</sup> evoked low freezing. During retrieval (Ret.), CS<sup>+</sup> but not CS<sup>-</sup> induced fear recovery ( $n = 21$  mice, Wilcoxon signed-rank test, CS<sup>-</sup> versus CS<sup>+</sup>;  $P < 0.001$ ). Error bars, mean  $\pm$  s.e.m. **c**, **d**, Left, raster plots and peristimulus time histograms (PSTHs) of CS<sup>+</sup>-evoked firing for INs (type 1 and 2) during Post FC (CS<sup>+</sup> 1–4, 108 trials). Right, mean z score of CS<sup>-</sup> and CS<sup>+</sup>-evoked responses of type 1 and type 2 INs during Post-FC, Ext. or Ret. sessions (CS<sup>-</sup> and CS<sup>+</sup> 1–4, 108 trials). Type 1 INs were excited ( $n = 68$ , 25 mice, paired  $t$ -test, CS<sup>-</sup> versus CS<sup>+</sup>,  $P < 0.001$ ), whereas type 2 INs were inhibited during CS<sup>+</sup> ( $n = 15$ , 8 mice, paired  $t$ -test, CS<sup>-</sup> versus CS<sup>+</sup>,  $P < 0.001$ ). Bins of 10 ms. **e**, **f**, Correlations between freezing during CS (Post-FC, Ext. or Ret. sessions, CS<sup>-</sup> and CS<sup>+</sup> 1–4) and CS-evoked firing (mean z score 0–150 ms post CS) for type 1 INs ( $n = 68$ , Pearson's  $r = 0.79$ ,  $P < 0.01$ ) and type 2 ( $n = 15$ , Pearson's  $r = -0.93$ ,  $P < 0.001$ ).

PVINs disinhibited PNs and produced freezing (Figs 2c and 3b). These data suggest that the increased activity of dmPFC PNs during fear expression results from a disinhibitory mechanism mediated by PVINs.

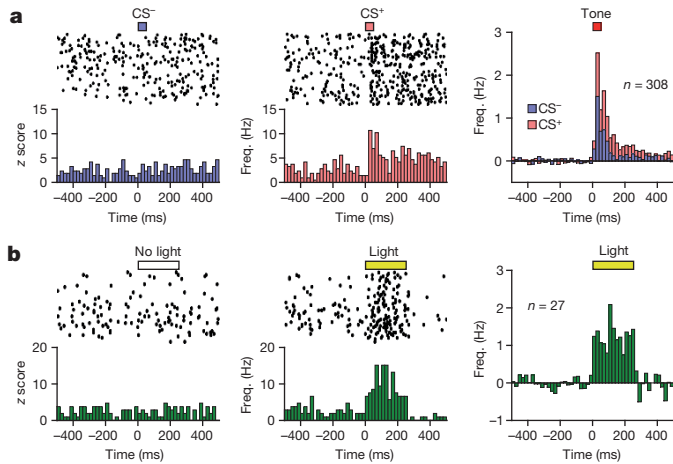
As PVINs have a key role in the genesis of cortical networks oscillations<sup>18,19</sup>, we investigated whether specific changes in dmPFC local field potentials (LFPs) were associated with different fear states. Although freezing periods were associated with a strong reduction of LFP theta power compared to non-freezing periods (Extended Data Fig. 7a), CS<sup>+</sup> but not CS<sup>-</sup> presentations were associated with a transient amplitude increase and a phase resetting of theta oscillations (Fig. 4a and Extended Data Fig. 7b). This analysis produced similar results when restricted to freezing and non-freezing periods during CS presentations (Extended Data Fig. 7c, d). This observation raises the question of whether dmPFC theta phase resetting during fear behaviour is mediated locally or imposed by a remote structure, such as the hippocampus. To address this question we locally injected muscimol to inactivate the medial septum, a structure that is involved in the genesis of hippocampal theta oscillations<sup>20</sup>. Inactivation of the medial septum reduced hippocampal



**Figure 2 | Prefrontal type 2 PVINs control fear expression.** **a**, Schematic of light inhibition of archaeorhodopsin (ArchT)-green fluorescent protein (GFP)-expressing PVINs (green) in dmPFC with an optic fibre coupled to the recording electrodes (yellow). Cg1, anterior cingulate cortex; PL, prelimbic area; IL, infralimbic area. **b**, PSTHs showing mean activity changes for type 1 (left,  $n = 9$ ) and type 2 INs (right,  $n = 5$ ) upon yellow light (yellow bars, 250 ms; 108 trials, 0.9 Hz). A Fisher exact statistical test revealed that the proportions of the two populations were significantly different ( $P = 0.033$ ). Bins of 10 ms. **c**, Protocol (left panel) and behaviour from PV-ires-Cre mice infected in dmPFC with GFP (control,  $n = 8$ ) or ArchT-GFP-expressing ( $n = 9$ ) floxed AAV viruses and submitted to yellow light. Before conditioning (Pre FC, middle panel) and after extinction (right panel), optogenetic inhibition of PVINs induced freezing (paired  $t$ -tests, Pre FC, GFP versus ArchT,  $***P < 0.001$ ; Extinction, GFP versus ArchT,  $***P < 0.001$ ; light-pulse duration, 250 ms; 108 trials, 0.9 Hz). **d**, Protocol (left panel) and behaviour from PV-ires-Cre mice infected with control GFP ( $n = 8$ ) or channelrhodopsin (ChR2)-enhanced yellow fluorescent protein (eYFP)-expressing ( $n = 6$ ) floxed AAV viruses in the dmPFC and submitted to blue light. Following conditioning (middle panel, Post FC), optogenetic activation of PVINs decreased freezing (Post FC, GFP versus ChR2, paired  $t$ -test,  $**P < 0.01$ ; light-pulse duration, 250 ms; 108 trials, 0.9 Hz). NS, not significant. Error bars, mean  $\pm$  s.e.m.

theta power, whereas it did not influence freezing and had no effect on dmPFC theta phase resetting evoked by CS<sup>+</sup> presentations (Extended Data Fig. 8).

Interestingly, we observed a strong correlation between CS<sup>+</sup>-evoked inhibition of PVINs and dmPFC theta phase resetting, suggesting that this phenomenon is gated by PVINs (Fig. 4b). In support of this hypothesis, optogenetic inhibition of PVINs reproduced theta phase resetting (Fig. 4c and Extended Data Fig. 9). Consistent with this, dmPFC theta resetting induced by CS<sup>+</sup> was blocked by optogenetic excitation of PVINs (Fig. 4d). Our results indicate that CS<sup>+</sup>-evoked inhibition of PVINs mediates theta phase resetting during fear expression, a phenomenon that might enhance synchronization and efficiency of dmPFC output neurons. To evaluate whether dmPFC theta phase resetting is associated with spiking synchronization among PNs during fear expression, we quantified the number of PNs displaying a significant firing increase during CS<sup>-</sup> and CS<sup>+</sup> presentations. Significantly more PNs were activated during CS<sup>+</sup> relative to CS<sup>-</sup> presentations (Fig. 5a). This activation was associated with a significant increase of coincident firing between pairs of PNs following CS<sup>+</sup> (Fig. 5b and Extended Data Fig. 10a). Furthermore, more PNs were significantly phase-locked to local theta oscillations during CS<sup>+</sup> relative to CS<sup>-</sup> presentations (Fig. 5c). Consistent with this, comparison of



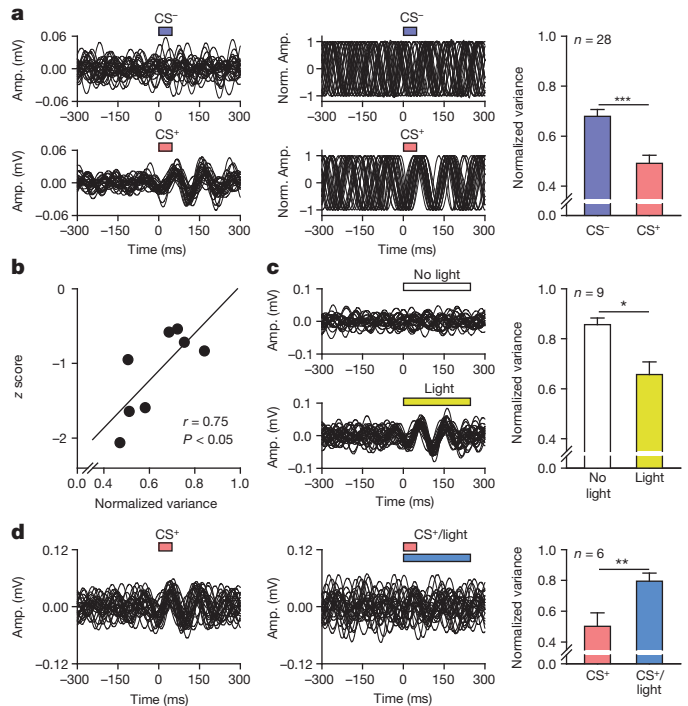
**Figure 3 | dmPFC PVINs disinhibit PNs during fear expression.** **a**, Raster plots and PSTHs illustrating CS<sup>-</sup> and CS<sup>+</sup>-evoked activities (left and middle panels, respectively) of a PN during Post FC (CS<sup>-</sup>/CS<sup>+</sup> 1–4, 108 trials). Right panel, mean PSTHs of PNs recorded during Post FC, Ext. or Ret. ( $n = 308$  neurons from 27 mice, CS<sup>-</sup> and CS<sup>+</sup> 1–4) showing a stronger and significant increase in response to CS<sup>+</sup> compared to CS<sup>-</sup>. **b**, Firing of a PN recorded in a mouse expressing ArchT in dmPFC PVINs at baseline (left panel, no light), and in response to yellow light (middle panel, light-pulse duration, 250 ms; 108 trials, 0.9 Hz). Right panel, mean PSTHs of all PNs displaying significant CS<sup>+</sup>-evoked excitation during Ext. and disinhibited during optogenetic inhibition of PVINs ( $n = 27/41$  PNs from 7 mice; light-pulse duration, 250 ms; 108 trials, 0.9 Hz). Bins of 20 ms.

the strength of theta phase locking, a measure of spiking synchronization, revealed a stronger tuning of dmPFC activity to local theta during CS<sup>+</sup> periods (Extended Data Fig. 10b). To evaluate whether enhancement of the spiking synchronization of PNs with local theta induced by CS<sup>+</sup> presentations was causally related to the inhibition of PVINs, we optogenetically manipulated PVINs and quantified PN theta phase locking. Our analysis revealed that light-induced inhibition of PVINs increased, whereas light-induced excitation of PVINs reduced PNs theta phase locking to dmPFC theta oscillations (Extended Data Fig. 10c, d).

To understand the dynamics of PNs synchronization during theta phase reset, the mean preferred phase of individual PNs was calculated during the first three theta cycles following CS<sup>+</sup> (Supplementary Methods). Relative to CS<sup>-</sup> presentations, CS<sup>+</sup>-induced firing of PNs occurred significantly more frequently around the peak of the oscillations, thereby creating precise temporal windows during which PNs were synchronized (Fig. 5d). Interestingly, similar to the CS<sup>+</sup> condition, artificial resetting of local theta oscillations, either by aligning the phase of individual LFPs during CS<sup>-</sup> presentations or by optogenetically inhibiting PVINs, produced synchronization of PNs firing around the peak of theta oscillations (Extended Data Fig. 10e, f). This observation suggests that the overall phase preference of PNs did not change between CS<sup>-</sup> or CS<sup>+</sup> conditions, but that PV-mediated theta phase resetting coordinated and sharpened synchronization among PNs.

Converging evidence indicates that dmPFC PNs target both the basolateral amygdala (BLA) and the periaqueductal grey (PAG), two structures involved in fear behaviour<sup>21–23</sup>. This raises the possibility that PNs may modulate fear expression through direct projections to the PAG and/or the BLA. To disentangle these possibilities, we antidromically activated dmPFC efferents using extracellular stimulation of BLA or PAG in anaesthetized mice, following completion of behaviour. These experiments revealed that PNs disinhibited during CS<sup>+</sup> presentations preferentially targeted the BLA (Fig. 5e, f and Extended Data Fig. 6d). These data indicate that theta phase resetting mediated by PVINs synchronizes PNs after CS<sup>+</sup> presentations and suggest that dmPFC PNs preferentially target the BLA to drive fear responses.

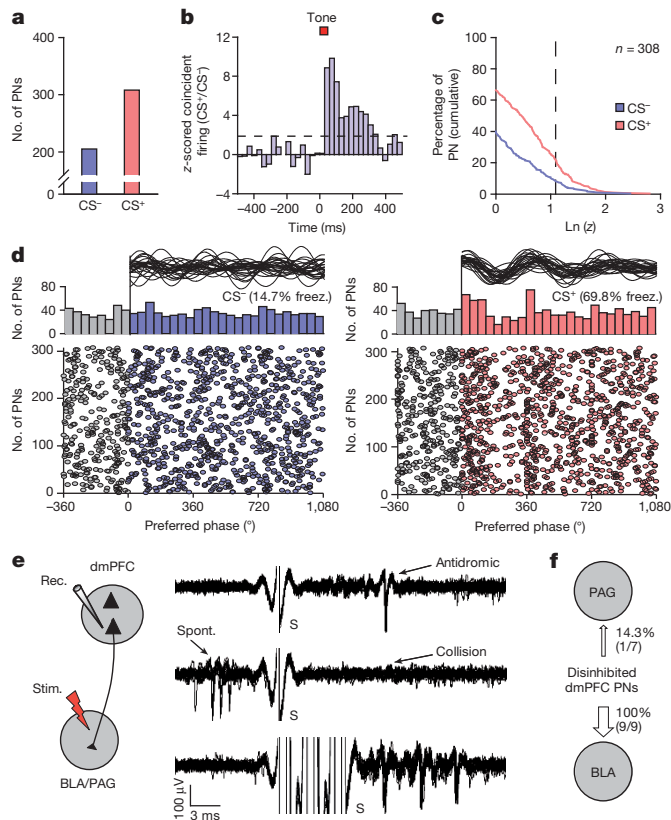
Using single-unit and LFP recordings combined with optogenetic manipulation of PVINs in mice, we have shown that a subpopulation



**Figure 4 | Inhibition of PVINs induces theta phase resetting.** **a**, dmPFC LFP trace amplitudes (Amp.) filtered in the 8–12-Hz range (left panel, 27 sweeps) and corresponding standardized trace amplitudes (middle panel) illustrating theta phase resetting induced by CS<sup>+</sup> but not CS<sup>-</sup> (Post-FC session, first CS<sup>-</sup> and first CS<sup>+</sup>, 27 sweeps). Right panel, time variance of the first theta peak following CS (Post FC, CS<sup>-</sup> and CS<sup>+</sup> 1–4,  $n = 28$  mice, CS<sup>-</sup> versus CS<sup>+</sup>, paired  $t$ -test, \*\*\* $P < 0.001$ ). **b**, Correlation between CS-evoked firing of type 2 INs (mean  $z$  score 0–150 ms post CS) and the time variance of the first theta peak following CS (Post FC, CS<sup>-</sup> and CS<sup>+</sup> 1–4,  $n = 15$  type 2 INs, Pearson's  $r = 0.75$ ,  $P < 0.05$ ). **c**, Left panel, dmPFC LFP traces recorded in a mouse expressing ArchT in PVINs in control conditions (top part; No light, 27 sweeps), and during optogenetic inhibition of PVINs (bottom part; light duration, 250 ms; 27 sweeps, 0.9 Hz). Right panel, time variance of the first theta peak following light ( $n = 9$  mice, No light versus Light, paired  $t$ -test, \* $P < 0.05$ ). **d**, Representative LFP traces recorded when CS<sup>+</sup> was presented alone (left panel, Post-FC session, first CS<sup>+</sup>) or paired with the optogenetic activation of PVINs (middle panel, Post-FC session, fifth CS<sup>+</sup>; light-pulse duration, 250 ms; 27 sweeps, 0.9 Hz). Right panel, time variance analyses of theta resetting during CS<sup>+</sup> or CS<sup>+</sup>/Light ( $n = 6$  mice, CS<sup>+</sup> versus CS<sup>+</sup>/Light, paired  $t$ -test, \*\* $P < 0.01$ ). Error bars, mean  $\pm$  s.e.m.

of PVINs organizes the spiking activity of dmPFC PNs during precise time windows, through phase resetting of local theta oscillations, to drive fear expression. Our data indicate that the fine regulation of dmPFC-BLA PNs by a subtype of PVINs is critical for the expression of fear behaviour. Our results demonstrate that inhibition of type 2 PVINs during CS<sup>+</sup> presentations is causally related to the expression of conditioned fear responses, and suggest that type 1 INs might inhibit type 2 INs. The origin of CS-mediated type 1 INs excitatory responses remains to be determined, but it is likely that they receive inputs from structures involved in the encoding or modulation of conditioned fear such as BLA or hippocampus<sup>24,25</sup>.

A key question is what mechanisms can account for our observation that inhibition of PVINs is necessary and sufficient for the expression of fear responses. Cortical PVINs are known to inhibit PNs through powerful perisomatic inhibition<sup>17</sup>. As a consequence, CS<sup>+</sup>-evoked inhibition in PVINs induced a strong disinhibition of PNs, a permissive mechanism that gated neuronal responses during fear expression. These results indicate that CS-evoked activity in dmPFC PNs during fear expression result in part from a disinhibitory mechanism. Notably, conditioned freezing was not entirely prevented by PVINs activation, indicating that some dmPFC PNs may escape inhibitory control, or



**Figure 5 | Synchronization of dmPFC PNs during fear expression.**

**a**, Number of significantly CS-activated PNs recorded during Post-FC, Ext. or Ret. sessions ( $CS^-$ ,  $n = 205$ ;  $CS^+$  1–4,  $n = 308$ ;  $CS^-$  versus  $CS^+$ , paired  $t$ -test,  $P < 0.001$ ). **b**, Normalized averaged ratio of changes in coincident activity between pairs of PNs induced by  $CS^-$  and  $CS^+$  (Post-FC, Ext. or Ret. sessions,  $n = 975$  pairs from 308 PNs). Dashed line, significant  $z$  score at  $P < 0.05$  level. Bins of 30 ms. **c**, Cumulative distribution of log-transformed Rayleigh's test  $Z$  of CS-responsive PNs ( $n = 308$ ; Post-FC, Ext. or Ret. sessions). Dashed line, significant theta phase locking threshold ( $\ln(Z) = 1.1$ ,  $P < 0.05$ ;  $CS^-$ ,  $n = 24$  neurons;  $CS^+$  1–4,  $n = 65$  neurons). **d**, Top panel, preferred theta phase distributions of PNs ( $n = 308$  neurons, 108 CS pips, Post-FC, Ext. or Ret. sessions) during theta cycles around  $CS^-$  (left part, blue bars, 14.7% freezing) and  $CS^+$  (right part, red bars, 69.8% freezing, bins of  $45^\circ$ ). Bottom panel, preferred theta phases of individual PNs. Example 8–12-Hz theta-filtered LFP traces during  $CS^-$  and  $CS^+$  are represented above for illustrative purposes. During  $CS^+$  but not  $CS^-$ , resetting of theta oscillations synchronizes the firing of PNs around the peaks of theta cycles (Rayleigh test for circular uniformity: first theta cycle post CS,  $CS^+$  versus  $CS^-$ ,  $P < 0.001$ ). **e**, Left panel, strategy used to identify connections between PNs and the BLA–PAG. Rec., recording electrode; S, stimulation artifact; Stim., stimulation electrode. Right panel, antidromic spikes recorded from a PN in response to BLA stimulations identified by their low temporal jitter (top trace, 10 trials), collisions with spontaneously occurring spikes (middle trace, 10 trials) and ability to follow high-frequency stimulation (bottom trace, 250 Hz, 10 trials). **f**, PNs exhibiting antidromic responses to BLA stimulations displayed  $CS^+$ -evoked excitation (9/9 neurons). Only a small fraction of PNs exhibiting antidromic responses to PAG stimulation displayed  $CS^+$ -evoked excitation (1/7 neurons, 14.3%). Thin arrow indicates that fewer neurons that project to the PAG are disinhibited; thick arrow indicates that more neurons that project to the BLA are disinhibited.

that other brain regions promote fear responses in concert with dmPFC.

Although fear behaviour was associated with a reduction in dmPFC theta-oscillation magnitude,  $CS^+$ -evoked inhibition of PVINs induced a robust and transient theta phase resetting spanning two to three theta cycles. Transient theta phase has been previously observed in cortical regions following electrical or sensory stimulations<sup>26–28</sup>. Our findings provide the first mechanistic explanation of phase resetting at the

cellular level and extend this phenomenon to the control of emotional behaviour. Functionally, we observed that theta phase resetting synchronized PNs around theta peaks without changing the preferred phases of individual PNs. This observation suggests that resetting of local theta oscillations, but not the preferred phases of individual PNs to the local LFP, is critically involved in the expression of fear responses. Thus, theta phase resetting represents a powerful mechanism for reliable fear expression because it creates an optimal temporal relationship that binds spiking activity with sensory information provided by CS. Ultimately, phase resetting of oscillations is a powerful mechanism that enhances the impact of input signals and enables transmission of information to downstream targets. Our data also show that reduction of rhythmic inhibition from PVINs paradoxically increases synchrony. Suppression of interference between two oscillators may account for this effect. Future work will be needed to identify the origin of dmPFC theta oscillations that are unmasked by PVINs inhibition.

Another question is how synchronized PNs can control fear expression. Previous findings suggest that putative dmPFC PNs displaying sustained or transient changes in their spiking activity promote fear expression through activation and synchronization of BLA neurons<sup>9,29,30</sup>. In line with these studies, our results demonstrate that PNs exhibiting  $CS^+$ -evoked synchronized firing during fear expression preferentially project to the BLA where they may target specific neuronal populations activated during fear behaviour<sup>23</sup>.

Finally, our findings suggest that persistent fear behaviour, which is at the core of psychiatric conditions such as anxiety disorders, may be finely regulated at the level of specific prefrontal inhibitory circuits.

## METHODS SUMMARY

Mice were submitted to a fear-conditioning paradigm in which the  $CS^+$  but not the  $CS^-$  was paired with a mild foot-shock (US). Extinction training was carried out over 2 days and mice were tested 1 week later for a retrieval session<sup>23</sup>. For optogenetic manipulations, PV-IRES-Cre mice received stereotaxic injections of AAV viruses encoding channelrhodopsin or archaeorhodopsin in the dmPFC. Bilateral activation of archaeorhodopsin or channelrhodopsin was performed using implanted optic fibres coupled to a laser beam. Inactivation of the medial septum was achieved using local pressure injection of fluorescently labelled muscimol. Individual neurons were recorded extracellularly and spikes were sorted by time-amplitude window discrimination and template matching as described<sup>23</sup>. CS-evoked responses were normalized to baseline activity using a  $z$ -score transformation. Antidromic and orthodromic spikes evoked by extracellular stimulations of the BLA or PAG were recorded in neurons isolated from behavioural sessions and recorded in urethane-anaesthetized mice, after completion of behaviour. *In vitro* whole-cell voltage and current-clamp recordings were performed using glass pipettes (4–6 M $\Omega$ ) filled with K-gluconate-based solutions.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 15 May; accepted 9 October 2013.**

**Published online 20 November 2013.**

1. Singer, W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49–65, 111–125 (1999).
2. Buzsáki, G. & Draguhn, A. Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004).
3. Womelsdorf, T. *et al.* Modulation of neuronal interactions through neuronal synchronization. *Science* **316**, 1609–1612 (2007).
4. Royer, S. *et al.* Control of timing, rate and bursts of hippocampal place cells by dendritic and somatic inhibition. *Nature Neurosci.* **15**, 769–775 (2012).
5. Cobb, S. R., Buhl, E. H., Halasy, K., Paulsen, O. & Somogyi, P. Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons. *Nature* **378**, 75–78 (1995).
6. Benchenane, K. *et al.* Coherent theta oscillations and reorganization of spike timing in the hippocampal–prefrontal network upon learning. *Neuron* **66**, 921–936 (2010).
7. Friedrich, R. W., Habermann, C. J. & Laurent, G. Multiplexing using synchrony in the zebrafish olfactory bulb. *Nature Neurosci.* **7**, 862–871 (2004).
8. Riehle, A., Grun, S., Diesmann, M. & Aertens, A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278**, 1950–1953 (1997).

9. Burgos-Robles, A., Vidal-Gonzalez, I. & Quirk, G. J. Sustained conditioned responses in prelimbic prefrontal neurons are correlated with fear expression and extinction failure. *J. Neurosci.* **29**, 8474–8482 (2009).
10. Tang, J. *et al.* Pavlovian fear memory induced by activation in the anterior cingulate cortex. *Mol. Pain* **1**, 6 (2005).
11. Vidal-Gonzalez, I., Vidal-Gonzalez, B., Rauch, S. L. & Quirk, G. J. Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. *Learn. Mem.* **13**, 728–733 (2006).
12. Corcoran, K. A. & Quirk, G. J. Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *J. Neurosci.* **27**, 840–844 (2007).
13. Pape, H. C. & Pare, D. Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. *Physiol. Rev.* **90**, 419–463 (2010).
14. Knapska, E. *et al.* Functional anatomy of neural circuits regulating fear and extinction. *Proc. Natl Acad. Sci. USA* **109**, 17093–17098 (2012).
15. LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
16. Hartwich, K., Pollak, T. & Klausberger, T. Distinct firing patterns of identified basket and dendrite-targeting interneurons in the prefrontal cortex during hippocampal theta and local spindle oscillations. *J. Neurosci.* **29**, 9563–9574 (2009).
17. Freund, T. F. & Katona, I. Perisomatic inhibition. *Neuron* **56**, 33–42 (2007).
18. Blatow, M. *et al.* A novel network of multipolar bursting interneurons generates theta frequency oscillations in neocortex. *Neuron* **38**, 805–817 (2003).
19. Losonczy, A., Zemelman, B. V., Vaziri, A. & Magee, J. C. Network mechanisms of theta related neuronal activity in hippocampal CA1 pyramidal neurons. *Nature Neurosci.* **13**, 967–972 (2010).
20. Yoder, R. M. & Pang, K. C. Involvement of GABAergic and cholinergic medial septal neurons in hippocampal theta rhythm. *Hippocampus* **15**, 381–392 (2005).
21. Gabbott, P. L., Warner, T. A., Jays, P. R., Salway, P. & Busby, S. J. Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. *J. Comp. Neurol.* **492**, 145–177 (2005).
22. Di Scala, G., Mana, M. J., Jacobs, W. J. & Phillips, A. G. Evidence of Pavlovian conditioned fear following electrical stimulation of the periaqueductal grey in the rat. *Physiol. Behav.* **40**, 55–63 (1987).
23. Herry, C. *et al.* Switching on and off fear by distinct neuronal circuits. *Nature* **454**, 600–606 (2008).
24. Sotres-Bayon, F., Sierra-Mercado, D., Pardilla-Delgado, E. & Quirk, G. J. Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. *Neuron* **76**, 804–812 (2012).
25. Tierney, P. L., Degenetais, E., Thierry, A. M., Glowinski, J. & Gioanni, Y. Influence of the hippocampus on interneurons of the rat prefrontal cortex. *Eur. J. Neurosci.* **20**, 514–524 (2004).
26. McCartney, H., Johnson, A. D., Weil, Z. M. & Givens, B. Theta reset produces optimal conditions for long-term potentiation. *Hippocampus* **14**, 684–687 (2004).
27. Buzsáki, G., Grastyán, E., Tveritskaya, I. N. & Czopf, J. Hippocampal evoked potentials and EEG changes during classical conditioning in the rat. *Electroencephalogr. Clin. Neurophysiol.* **47**, 64–74 (1979).
28. Rizzuto, D. S. *et al.* Reset of human neocortical oscillations during a working memory task. *Proc. Natl Acad. Sci. USA* **100**, 7931–7936 (2003).
29. Livneh, U. & Paz, R. Amygdala-prefrontal synchronization underlies resistance to extinction of aversive memories. *Neuron* **75**, 133–142 (2012).
30. Chang, C. H., Berke, J. D. & Maren, S. Single-unit activity in the medial prefrontal cortex during immediate and delayed extinction of fear in rats. *PLoS ONE* **5**, e11971 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank members of the Herry laboratory, K. Benchenane and D. Dupret for comments on the manuscript, K. Deisseroth and E. Boyden for generously sharing material, J. Babelo, S. Wolff and P. Tovote for technical and computational assistance, the Bordeaux Imaging center of the University of Bordeaux, and C. Poujol and S. Marais for technical assistance with microscopy. This work was supported by grants from the French National Research Agency (ANR-2010-BLAN-1442-01; ANR-10-EQPX-08 OPTOPATH), the European Research Council (ERC) under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement no. 281168, a Fonds AXA pour la recherche doctoral fellowship (J.C.) and the Conseil Régional d'Aquitaine. T.C.M.B. is a fellow of Ecole de l'Inserm Liliane Bettencourt-MD-PhD program, France.

**Author Contributions** J.C., F.C., R.R.R., N.K., C.G.-C., H.W., A.A., J.B. and T.C.M.B. performed the experiments and analysed the data. J.C. and C.H. designed the experiments and wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.H. ([cyril.herry@inserm.fr](mailto:cyril.herry@inserm.fr)).

## METHODS

**Animals.** Male C57BL6/J mice (3 months old, Janvier) and PV-IRES-Cre mice (3 months old, Jackson Laboratory, B6;129P2-Pvalb<sup>tm1(Cre)Arbr/J</sup>) were individually housed for at least 7 days before all experiments, under a 12-h light–dark cycle, and provided with food and water *ad libitum*. All procedures were performed in accordance with standard ethical guidelines (European Communities Directive 86/60-EEC) and were approved by the committee on Animal Health and Care of Institut National de la Santé et de la Recherche Médicale and French Ministry of Agriculture and Forestry (authorization A3312001).

**Behaviour.** Fear conditioning and extinction took place in two different contexts (context A and B). The conditioning and extinction boxes were cleaned with 70% ethanol and 1% acetic acid before and after each session, respectively. To score freezing behaviour, an automated infrared beam detection system located on the bottom of the experimental chambers was used (Coulbourn Instruments). The animals were considered to be freezing if no movement was detected for 2 s. On day 1, C57BL6/J mice were submitted to an habituation session in context A, in which they received four presentations of the CS<sup>+</sup> and of the CS<sup>−</sup> (total CS duration, 30 s; consisting of 50-ms pips at 0.9 Hz repeated 27 times, 2 ms rise and fall; pip frequency, 7.5 kHz or white-noise, 80 dB sound pressure level). Discriminative fear conditioning was performed on the same day by pairing the CS<sup>+</sup> with a US (1-s foot-shock, 0.6 mA, 5 CS<sup>+</sup>–US pairings; inter-trial intervals, 20–180 s). The onset of the US coincided with the offset of the CS<sup>+</sup>. The CS<sup>−</sup> was presented after each CS<sup>+</sup>–US association but was never reinforced (five CS<sup>−</sup> presentations; inter-trial intervals, 20–180 s). The frequencies used for CS<sup>+</sup> and CS<sup>−</sup> were counterbalanced across animals. On day 2 and day 3, conditioned mice were submitted to extinction training (post-fear-conditioning and extinction sessions) in context B during which they received 4 and 12 presentations of the CS<sup>−</sup> and CS<sup>+</sup>, respectively. Retrieval of fear was tested 7 days later in context B, with 4 presentations of the CS<sup>−</sup> and the CS<sup>+</sup>. Four distinct behavioural experiments were performed to collect the entire data set.

For optogenetic experiments using archaeorhodopsin, PV-IRES-Cre mice were submitted on day 1 to a pre-fear-conditioning session in context A during which they received yellow light stimulations (250-ms pulses repeated at 0.9 Hz during 2 min). Fear conditioning was performed on day 2 in context A, by pairing the CS<sup>+</sup> with the US (1-s foot-shock, 0.6 mA, 5 CS<sup>+</sup>/US pairings; inter-trial interval, 20–180 s). On day 2 and day 3, conditioned mice were submitted to extinction training (post-fear-conditioning and extinction sessions) in context B during which they received 12 presentations of the CS<sup>+</sup>. At the end of the last extinction session they received an additional four presentations of the CS<sup>+</sup> coupled to yellow light stimulations (each CS<sup>+</sup> pip was paired with a 250-ms light pulse). For optogenetic experiments using archaeorhodopsin, two distinct behavioural experiments were performed to collect the entire data set. For optogenetic experiments using channelrhodopsin, PV-IRES-Cre mice were submitted on day 1 to the same fear conditioning protocol as above. A post-conditioning test was performed on day 2 in context B and consisted of four presentations of the CS<sup>+</sup> alone followed by four presentations of the CS<sup>+</sup> coupled to blue light stimulations (each CS<sup>+</sup> pip was paired with a 250-ms light pulse). On day 3, mice were submitted to a second test in context B (Test) in which they received four presentations of the CS<sup>+</sup>. For optogenetic experiments using channelrhodopsin, two distinct behavioural experiments were performed to collect the entire data set.

For the place-avoidance experiment, we used an apparatus composed of two plexiglas compartments (20 × 10 cm each) connected by an alleyway. The two compartments differed tactilely (smooth plastic versus metal bars) and visually (grey plexiglas with red horizontal stripes or grey plexiglas). The time spent in each compartment was automatically recorded by an infrared beam detection system located on the bottom of the apparatus (Imetronic). On day 1, mice were allowed to explore freely the entire apparatus during a 15-min pre-exposure. Following pre-exposure, the compartment in which the mice spent the most time was designated as the most visited compartment. On day 2, mice were submitted to a 15-min test session during which light pulses (250-ms pulse width, repeated at 0.9 Hz) were delivered while the animals occupied the most visited compartment, but not when they occupied the less-visited compartment. The automated infrared beam sensors detected when the animal fully entered and exited the most visited compartment on day 1. The laser was automatically turned on for the period of time in which the animal stayed in the most visited compartment. The laser was turned on only when the animal fully entered the most visited compartment, not before the entrance. For place avoidance experiments, two distinct behavioural experiments were performed to collect the entire data set.

For pharmacological experiments, C57BL6/J mice were submitted to a fear conditioning paradigm consisting of CS<sup>+</sup> and US pairings in context A as described above. On days 2, 3 and 4, conditioned mice were tested in context B during which they received four presentations of the CS<sup>+</sup> before muscimol injections (Day 2, Test pre-MUS), 5 min after muscimol injections (Day 3, Test MUS), and 24 hrs

following muscimol injections (Day 4, Test post-MUS). For pharmacological experiments, two distinct behavioural experiments were performed to collect the entire data set.

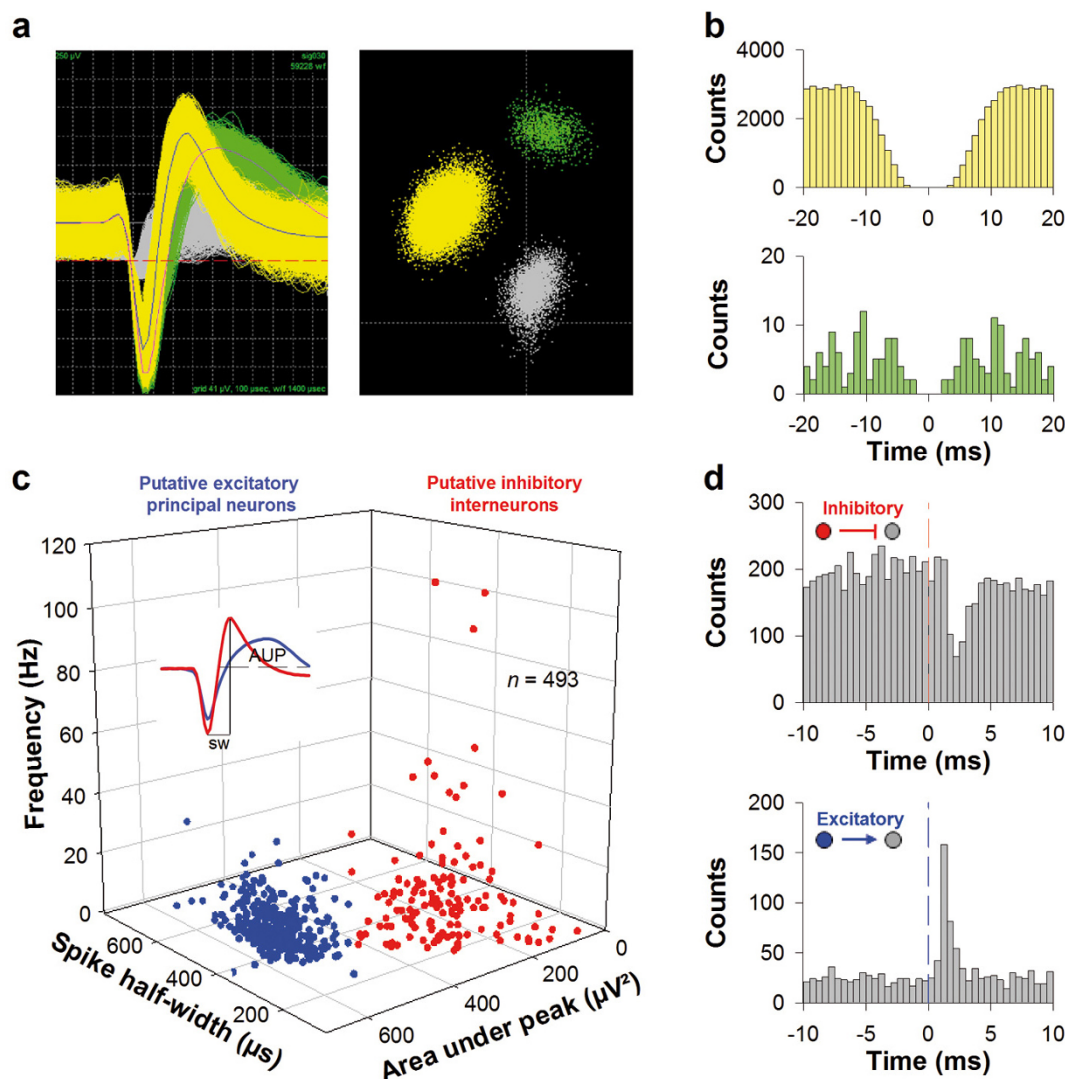
**Surgery and recordings.** Mice were anaesthetized with isoflurane (induction 3%, maintenance 1.5%) in O<sub>2</sub>. Body temperature was maintained at 37 °C with a temperature controller system (FHC). Mice were secured in a stereotaxic frame and unilaterally implanted in the left dorsomedial prefrontal cortex (dmPFC) with a multi-wire electrode array aimed at the following coordinates: 2 mm anterior to bregma; 0.3 mm lateral to the midline; and 0.8 to 1.4 mm ventral to the cortical surface. A subset of animals (*n* = 10) were also implanted in the dorsal hippocampus (dHip) at the following coordinates: 2 mm posterior to bregma; 1.2 mm lateral to midline; and 1.2 to 1.4 mm ventral to the cortical surface. The electrodes consisted of 16 individually insulated nichrome wires (13 µm inner diameter, impedance 30–100 KΩ; Kanthal) contained in a 26-gauge stainless-steel guide cannula. The wires were attached to an 18-pin connector (Omnetics). For mice that received dmPFC and dHip multi-wire implants, two connectors were used. All implants were secured using Super-Bond cement (Sun Medical). After surgery mice were allowed to recover for 7 days and were habituated to handling. Analgesia was applied before, and 1 day after surgery (Metacam, Boehringer). Electrodes were connected to a headstage (Plexon) containing sixteen unity-gain operational amplifiers. The headstage was connected to a 16-channel preamplifier (gain 100× bandpass filter from 150 Hz to 9 kHz for unit activity and from 0.7 Hz to 170 Hz for field potentials, Plexon). Spiking activity was digitized at 40 kHz and bandpass filtered from 250 Hz to 8 kHz, and isolated by time-amplitude window discrimination and template matching using a Multichannel Acquisition Processor system (Plexon). At the conclusion of the experiment, recording sites were marked with electrolytic lesions before perfusion, and electrode tips locations were reconstructed with standard histological techniques.

**Single-unit analyses.** Single-unit spike sorting was performed using Off-Line Spike Sorter (OFSS, Plexon) for all behavioural sessions. Principal-component scores were calculated for unsorted waveforms and plotted in a three-dimensional principal-component space; clusters containing similar valid waveforms were manually defined. A group of waveforms were considered to be generated from a single neuron if the waveforms formed a discrete, isolated, cluster in the principal-component space and did not contain a refractory period less than 1 ms, as assessed using auto-correlogram analyses. To avoid analysis of the same neuron recorded on different channels, we computed cross-correlation histograms. If a target neuron presented a peak of activity at a time that the reference neuron fired, only one of the two neurons was considered for further analysis. After fear conditioning, if the same neuron was sequentially recorded during different behavioural sessions, we considered only the first behavioural session in which it was recorded. To separate putative inhibitory interneurons (INs) from putative excitatory principal neurons (PNs) we used an unsupervised cluster algorithm based on Ward's method. In brief, the Euclidian distance was calculated between all neuron pairs based on the three-dimensional space defined by each neuron's average half-spike width (measured from trough to peak), the firing rate and the area under the hyperpolarization phase of the spike. An iterative agglomerative procedure was then used to combine neurons into groups based on the matrix of distances such that the total number of groups was reduced to give the smallest possible increase within-group sum of square deviation. To assess the significance of cross-correlogram analyses performed between pairs of recorded neurons, a mean firing rate with 95% confidence limits of the target neuron was calculated. Significant short-latency inhibitory or excitatory interactions were retained if the number of action potentials of the target neuron was inferior or superior to the 95% confidence limits, respectively. Moreover, to show that cross-correlations were not simply occurring by chance or were due to CS presentations, we performed two controls. First, the spike train of the neuron was shuffled 100 times and a shuffled cross-correlogram was computed. Absence of short-latency interaction in the shuffled cross-correlogram was indicative that the cross-correlations were not due to chance. Second, to control that short-latency interactions were not artificially induced by stimulus presentations, we computed a shift predictor and subtracted it from the original cross-correlogram. Persistence of short-latency cross-correlations indicates that the neuronal interactions were not due to CS presentations. CS or light-induced neural activity of recorded neurons was calculated by comparing the firing rate after stimulus onset with the firing rate recorded during the 500 ms before stimulus onset (bin size of 10 ms) using a *z*-score transformation. *z*-score values were calculated by subtracting the average baseline firing rate established over the 500 ms preceding stimulus onset from individual raw values and by dividing the difference by the baseline standard deviation. Only CS<sup>+</sup> responsive neurons (at least one significant positive or negative *z*-score bin (*z*-score > ± 1.67, *P* < 0.05) within 100 ms following CS onset) were considered for further analysis. For statistical analysis, *z*-score comparisons were performed using the average *z*-score value calculated during the 150 ms after CS onset.

To identify the main firing patterns among INs, we used an unbiased principal component analysis (PCA) based on the neuronal activity evoked by CS<sup>+</sup> presentations (z-score 500 ms before and after CS<sup>+</sup> presentations, CS<sup>+</sup> presentations 1–4 in post-fear-conditioning, extinction and retrieval sessions, each CS<sup>+</sup> consisting of 27 individual sound pips; bin size of 10 ms). Only the first principal component was considered (PC1) because it explained most of the variance of our data set. Type 1 and type 2 interneurons were defined as correlated and inversely correlated, respectively, with PC1 at the  $P < 0.001$  significance level. Co-firing between recorded PN pairs was established by quantifying the number of non-overlapping 30-ms time windows following CS<sup>+</sup> presentations during which co-firing events occurred (each pip presentation, CS<sup>+</sup> presentations 1–4, 108 pips, post-fear-conditioning, extinction or retrieval sessions). We then calculated a ratio of coincident firing by dividing the number of co-firing occurrences during CS<sup>+</sup> presentations by those obtained during CS<sup>−</sup> presentations. This coincident firing ratio was normalized to the pre-CS period (500 ms pre CS) using a z-score transformation. To control that the changes in coincident firing between CS<sup>+</sup> and CS<sup>−</sup> conditions were not due to an increase in PN firing rate during CS<sup>+</sup> presentations, the same analysis was performed but this time the number of co-firing events in each 30-ms time window was normalized by the total number of spikes of the two neurons in this particular time window. Statistical analyses were performed using paired Student's *t*-tests post hoc comparisons at the  $P < 0.05$  level of significance unless indicated otherwise. Results are presented as mean  $\pm$  s.e.m.

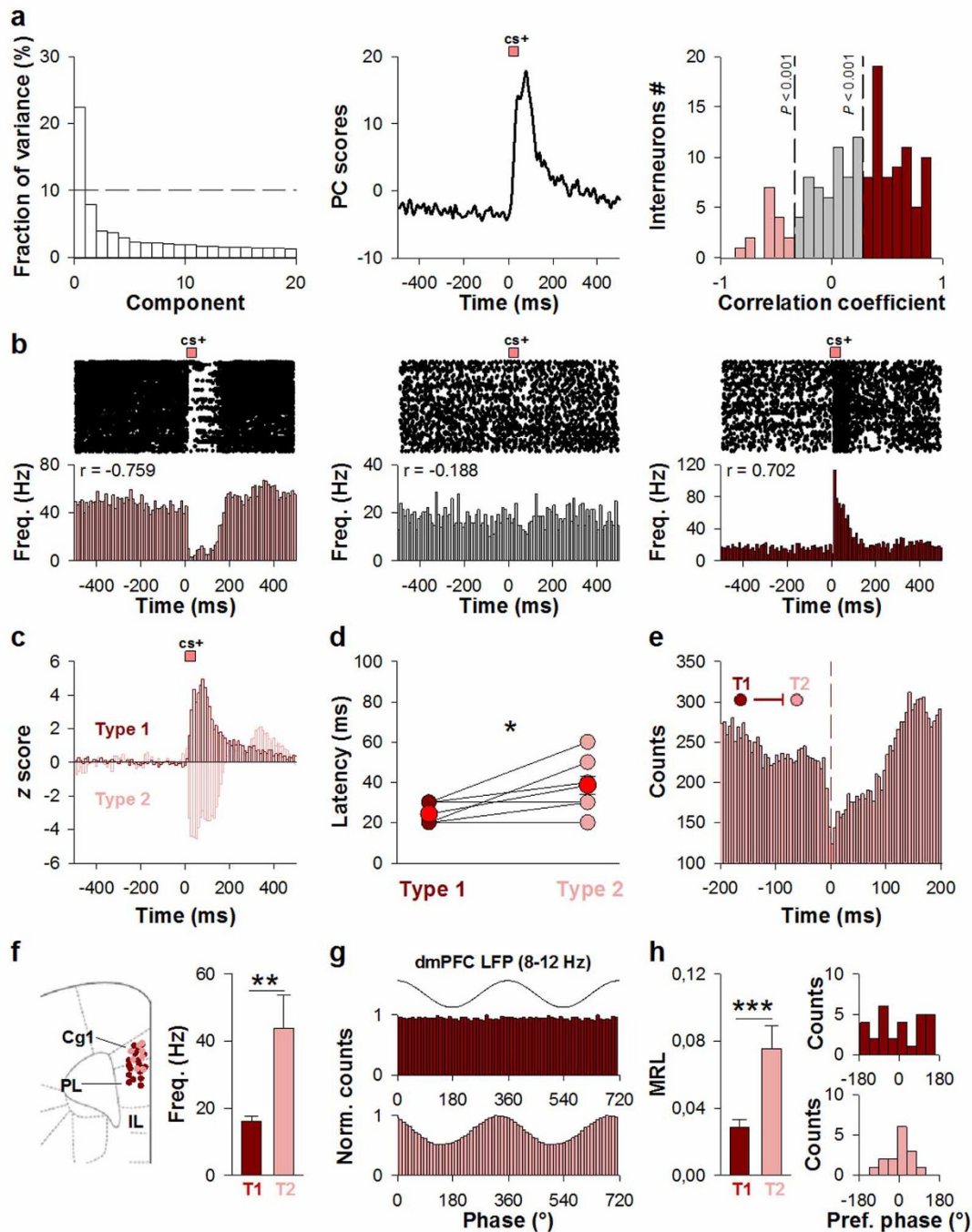
**Statistical analyses.** For each statistical analysis provided in the manuscript, the Kolmogorov–Smirnov normality test was first performed on the data to determine whether parametric or non-parametric tests were required. Two different approaches were used to calculate the sample size. For studies in which we had sufficient information on response variables, power analyses were carried out to determine

the number of mice needed. For studies in which the behavioural effect of the manipulation could not be pre-specified, such as optogenetic experiments, we used a sequential stopping rule (SSR). In essence this method enables null-hypothesis tests to be used in sequential stages, by analysing the data at several experimental points using *t*-tests. Usually the experiment started by testing only a few animals and if the *P* value was below 0.05, the investigator declared the effect significant and stopped testing. If the *P* value was greater than 0.36, the investigator stopped the experiment and retained the null hypothesis. For sample-size estimation using power analyses, we used an on-line power analysis calculator (G\*power3). For each analysis, sample size was determined using a power  $> 0.9$  and alpha error = 0.05. All tests were two sided. Power analyses were computed for matched pairs (differential conditioning protocol in which we used an internal control (Fig. 1) and pharmacological experiments (Extended Data Fig. 8)). In our behavioural experiments, a critical parameter is freezing level, and the numerical endpoint typically ranges between 50 and 70% freezing for CS<sup>+</sup> presentations immediately following auditory fear conditioning and between 10 and 30% freezing for CS<sup>−</sup> presentations. A minimum biologically significant difference in the mean values between CS<sup>−</sup> and CS<sup>+</sup> conditions (Fig. 1), or between CS<sup>+</sup> presentations before and after pharmacological treatment (Extended Data Fig. 8) is 1.5-fold. If we assume a standard deviation of 1.5 for a mean value of 60% freezing for CS<sup>+</sup> and 20% freezing for CS<sup>−</sup> or CS<sup>+</sup> after pharmacological treatment (which are realistic numbers), then a minimal  $n = 6$  (paired *t*-test) or  $n = 8$  (Wilcoxon signed-rank test) is needed to reject the null hypothesis with 90% probability. Sample size determination using SSRs analyses were used for optogenetic experiments in which it was not possible to determine a priori the effect of the optical manipulation. We used *P* values of 0.05 and 0.36 for lower and upper criterion. Using this strategy we ended up with an  $n$  comprising between 6 and 13 animals per group.



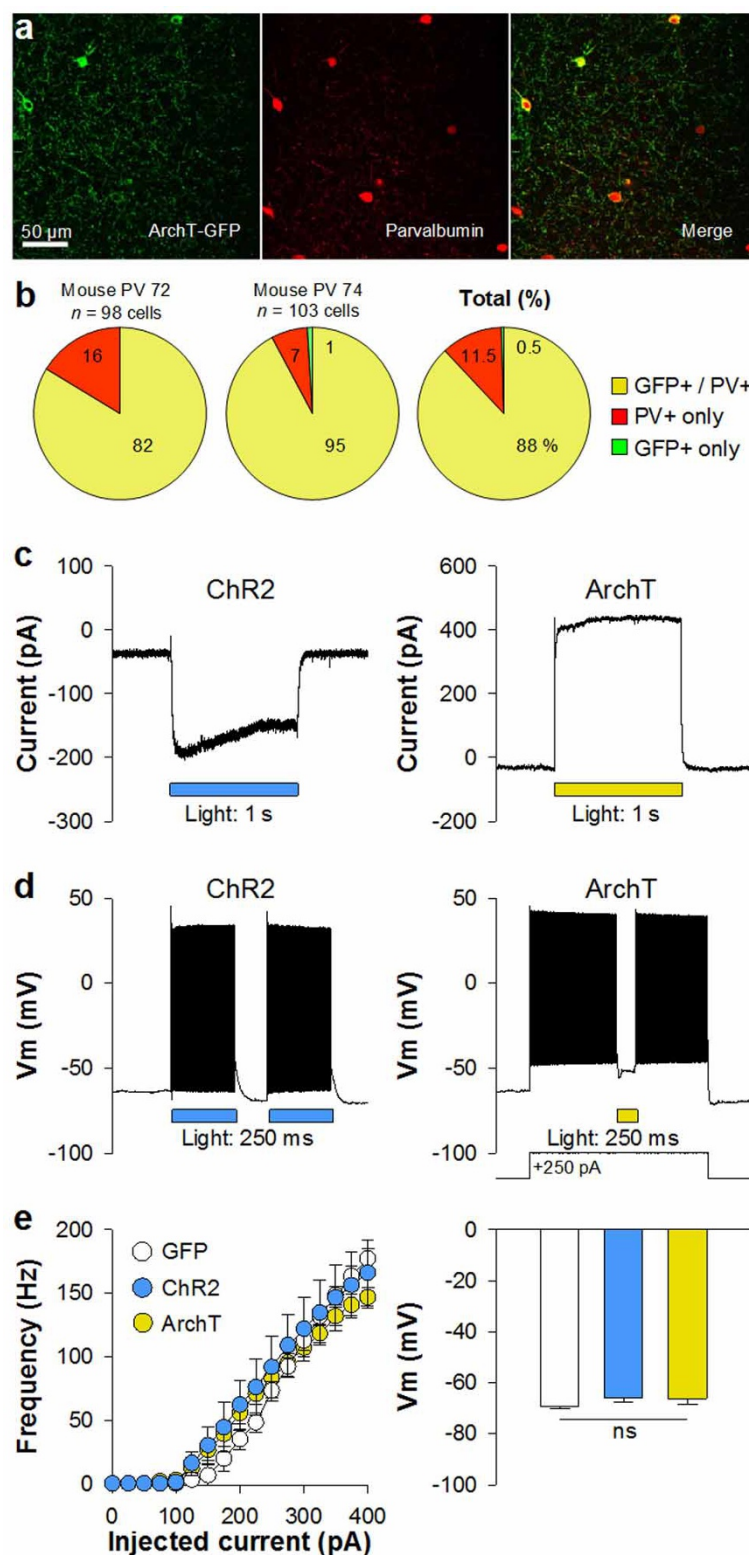
**Extended Data Figure 1 | Separation of putative principal neurons and putative interneurons.** **a**, Left panel, superimposed waveforms recorded from two different units. Right panel, spikes originating from individual units were sorted using three-dimensional principal-component analysis. **b**, Corresponding auto-correlograms, colour-coded as in **a**, displaying clear refractory periods. **c**, Among the population of dmPFC neurons displaying significant excitatory or inhibitory CS<sup>+</sup>-evoked responses ( $n = 493$ ), 71.2% were classified as putative principal neurons (PNs, blue circles,  $n = 351$ ) and 28.8% as putative interneurons (INs, red circles,  $n = 142$ ) using an unbiased unsupervised cluster-separation algorithm based on three electrophysiological properties: firing frequency, spike half-width and spike area under waveform (AUP) peak. Inset, average waveform of a representative PN and IN illustrating

the methodology used to quantify spike width (SW) and the spike segment used to calculate the AUP. **d**, Top panel, representative cross-correlogram performed between a putative inhibitory IN and a non-identified neuron showing a short-latency, presumably monosynaptic, inhibitory interaction (7 pairs identified among putative INs, no inhibitory interaction among putative PNs). Bottom panel, representative cross-correlogram between a putative PN and a non-identified neuron showing a short-latency, possibly monosynaptic, excitatory interaction (20 pairs identified among PNs, no excitatory interaction from putative INs). Reference events correspond to the spikes of the pre-synaptic neuron (dashed line at time 0, bins of 0.5 ms). Grey circles represent neurons that were not tone-responsive.



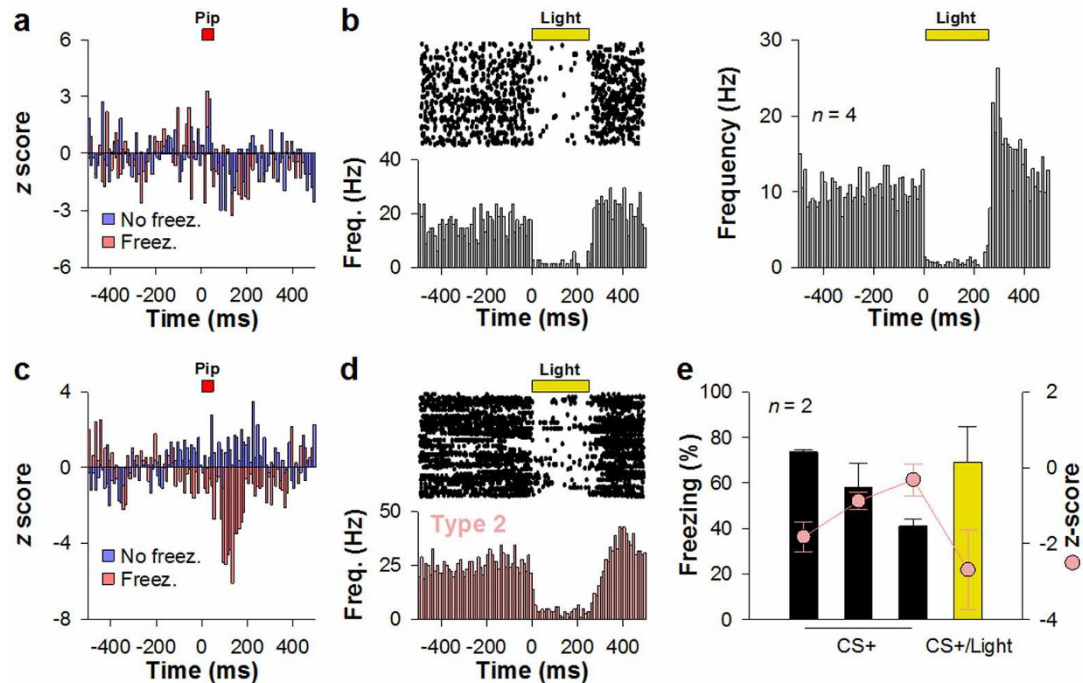
**Extended Data Figure 2 | CS<sup>+</sup>-evoked firing patterns and inhibitory interactions of putative INs.** **a**, Left panel, distribution of the fraction of variance for the 20 first principal components (PCs) obtained with principal component analysis (PCA). PC1, which accounted for more than 20% of variance of the data set, was used for the analysis. Middle panel, first principal-component coefficients representing the main firing pattern evoked by CS<sup>+</sup> (CS<sup>+</sup> onset at time 0) of the IN data set. Right panel, distribution of dmPFC IN correlation coefficients with PC1. The dashed lines indicate the levels of significance ( $P < 0.001$ ). Among the 142 INs, 83 (58.5%) displayed a significant positive ( $n = 68$ , 48%, dark red bars) or negative ( $n = 15$ , 10.6%, light red bars) correlation with PC1, whereas 41.5% INs ( $n = 59$ , grey bars) did not. **b**, Raster plots and PSTH of individual INs negatively correlated (left part, type 2 IN), not correlated (middle part) or positively correlated (right part, type 1 IN) with PC1. Type 1 INs were excited, whereas type 2 INs were inhibited by CS<sup>+</sup>. Bins of 10 ms. **c**, PSTH of all type 1 ( $n = 68$ ) and type 2 ( $n = 15$ ) INs illustrating the CS<sup>+</sup>-evoked responses (Post-FC, Ext. or Ret. Sessions, CS<sup>+</sup> 1–4). Bins of 10 ms. **d**, Individual (type 1 INs, dark red dots; type 2 INs, light red dots) and averaged (red dots) latencies of the first significant time bin ( $z$  score  $< -1.65$  or

$> +1.65$ ) following CS<sup>+</sup> for type 1 and type 2 INs recorded simultaneously ( $n = 7$  pairs recorded in 5 mice). CS<sup>+</sup>-evoked excitation in type 1 INs preceded CS<sup>+</sup>-evoked inhibition in type 2 INs (mean latency: type 1,  $24.3 \pm 2$  ms; type 2,  $38.6 \pm 4.6$  ms; paired  $t$ -test,  $*P < 0.05$ ). Error bars, mean  $\pm$  s.e.m. **e**, Cross-correlation analysis performed between a type 1 and a type 2 IN recorded simultaneously outside CS. The cross-correlogram shows a short latency, potentially monosynaptic, inhibitory interaction. Reference event, spikes of the type 1 IN (dashed line at time 0). Bins of 5 ms. **f**, Locations of recording sites and mean firing frequencies of type 1 (T1,  $n = 68$ ) and type 2 (T2,  $n = 15$ ) INs (Mann–Whitney test,  $**P < 0.01$ ; Cg1, anterior cingulate cortex; PL, prelimbic area; IL, infralimbic area). **g**, Firing modulation of representative type 1 and type 2 INs with dmPFC theta oscillations filtered in the 8–12-Hz range (12-min recordings). Bins of 10°. **h**, Mean strength of firing synchronization to local theta oscillations as measured with the mean resultant length (MRL) vector (left panel, Mann–Whitney test, type 1 versus type 2,  $***P < 0.001$ ) and distribution of the preferred phases (right panels) for type 1 and type 2 INs significantly phase-locked to theta oscillations (type 1,  $n = 29/68$ ; type 2,  $n = 15/15$ ).



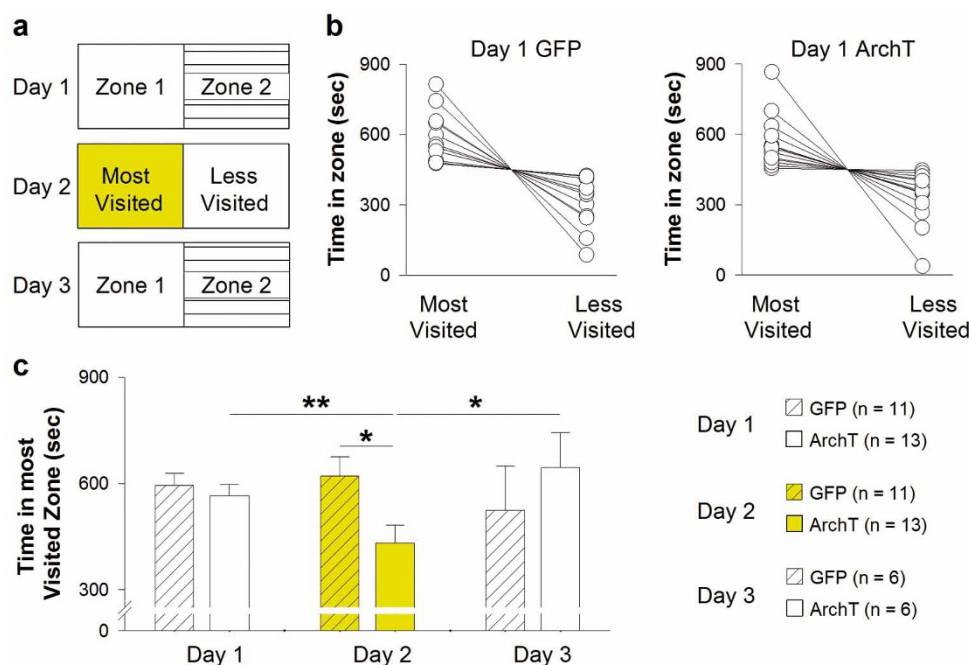
**Extended Data Figure 3 | Anatomical characterization of AAV-mediated ArchT-GFP expression in PV-IRES-Cre transgenic mice and electrophysiological characteristics of ArchT, ChR2 and GFP PV-IRES-Cre-infected PV neurons.** **a**, Representative confocal micrographs used for PV and GFP co-localization assessment. Left panel, ArchT-GFP labelled with anti-GFP Alexa 488; middle panel, PV immunofluorescence; right panel, merge. Single optical slices, in the same focal plane. **b**, Quantitative analysis of viral infection specificity and efficacy. Pie charts show the numbers of neurons positive for GFP and/or PV in two mice (left and middle charts) and averaged proportions (right chart). **c**, Representative ChR2- (left) and ArchT-evoked (right) currents recorded from PVINs with somata located in layer 2/3 of the

dmPFC (1 s light stimulation). **d**, Representative optically evoked action potential firing and inhibition of PVINs expressing ChR2 (left, 500-ms blue light pulses) or ArchT, respectively (right, 250-ms yellow light pulse during a 250-pA current pulse injection). **e**, Left panel, changes in firing frequency of PVINs expressing GFP (white dots,  $n = 7$ ), ChR2 (blue dots,  $n = 5$ ) or ArchT (yellow dots,  $n = 8$ ) upon injection of increasing current pulses (current pulses range, 0–400 pA). No significant differences were observed between groups. Right panel, resting membrane potentials of INs expressing GFP (white bar,  $n = 7$ ), ChR2 (blue bar,  $n = 5$ ) or ArchT (yellow bar,  $n = 8$ ). No significant differences were observed between groups (unpaired  $t$ -tests).



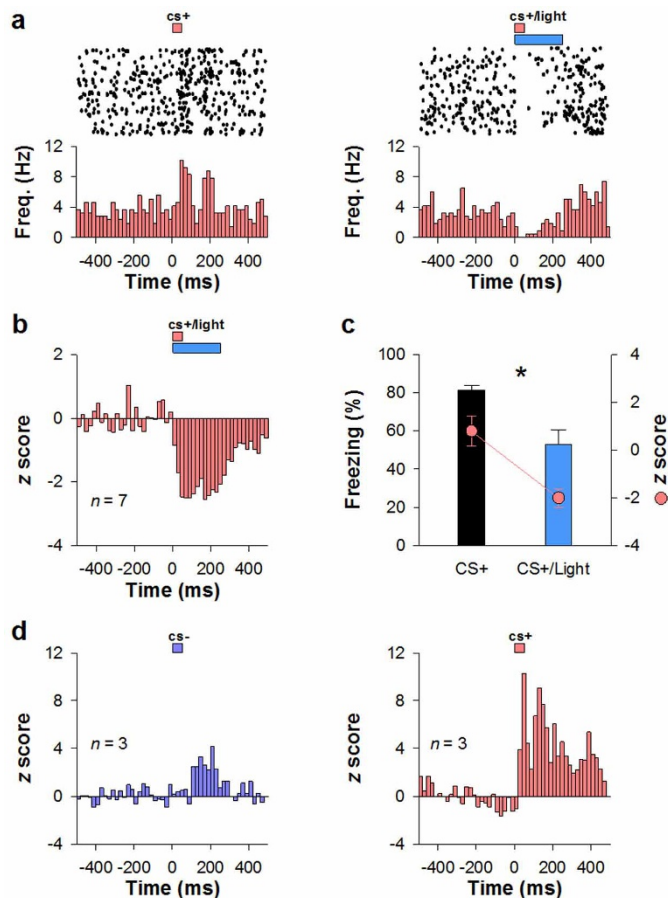
**Extended Data Figure 4 | Type 2 PVINs mediate conditioned fear responses.** **a**, z-score transformation of CS<sup>+</sup>-evoked firing of a non-type-2 IN for sound pips outside (No freez.) or inside (Freez.) freezing periods during the extinction session (CS<sup>+</sup> 1–12; No freez., 141 pips; Freez., 156 pips). This neuron was not classified as a type 1 or type 2 IN. **b**, Left panel, raster plot illustrating optogenetic identification of the same non-type-2 IN as ArchT-expressing (that is, PV-expressing). Right panel, mean z-score transformation of all non-type-2 INs identified as PV-expressing INs ( $n = 4$ ; light-pulse duration, 250 ms; 108 stimulation trials). **c**, z-score transformation of CS<sup>+</sup>-evoked firing of a type 2 IN for No freez. and Freez. periods during

the extinction session (CS<sup>+</sup> 1–12; No freez., 141 pips; Freez., 156 pips). **d**, Raster plot illustrating optogenetic identification of the same type 2 IN as ArchT-expressing (that is, PV-expressing) (light-pulse duration, 250 ms; 108 stimulation trials). **e**, CS<sup>+</sup>-evoked changes in firing rate in two type 2 PVINs identified with optogenetic, and corresponding freezing scores of the two mice in which they were recorded (dots, mean z-score 150 ms post CS; bars, blocks of 4 CS<sup>+</sup> presentation each, both during the second extinction session; light-pulse duration, 250 ms; 108 stimulation trials). Light-induced inhibition of PV, including type 2 INs, reinstated freezing behaviour. Error bars indicate mean  $\pm$  s.e.m.

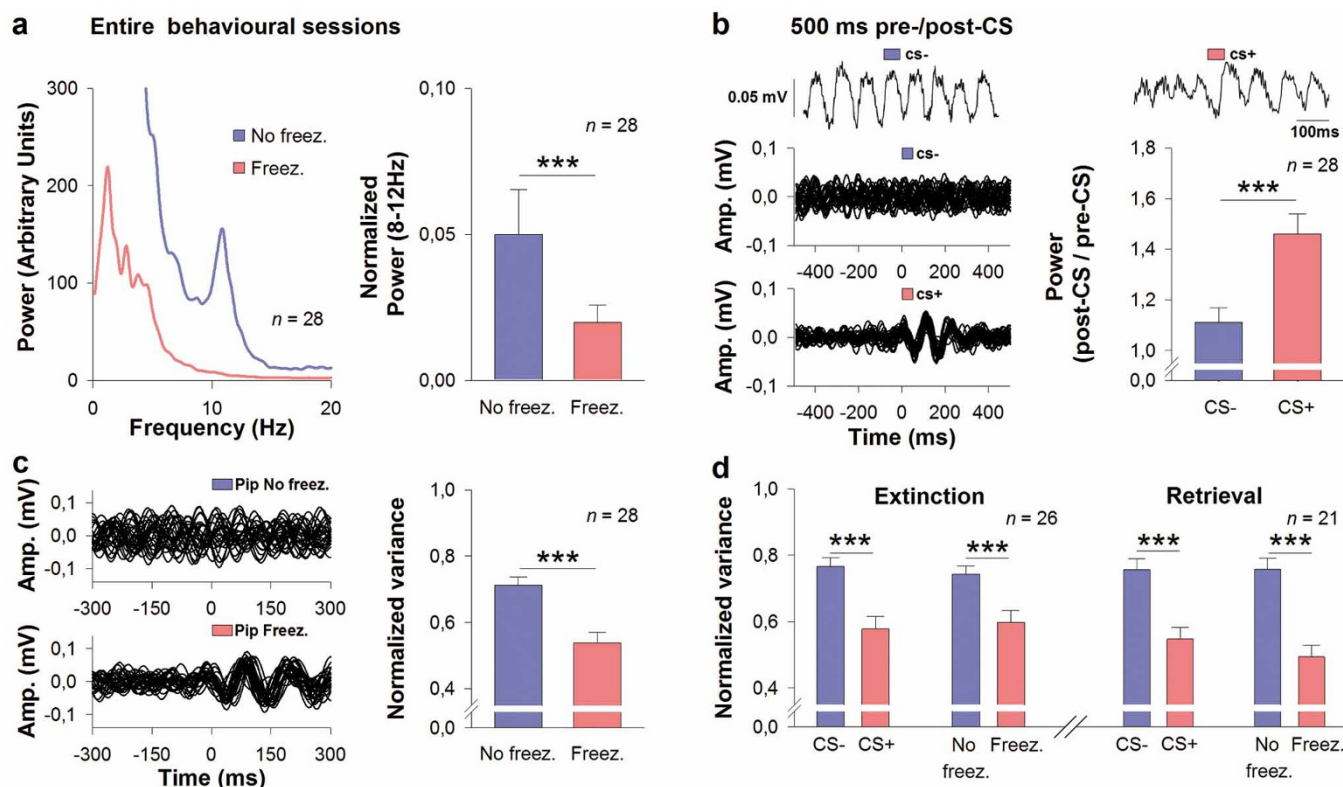


**Extended Data Figure 5 | Optogenetic inhibition of prefrontal PVINs induces place aversion.** **a**, On day 1, GFP- and ArchT-infected mice ( $n = 11$  and  $13$ , respectively) were exposed to a two-compartment place aversion apparatus during 15 min. Following pre-exposure, the most visited compartment was selected for each animal. On day 2, systematic yellow-light-induced inhibition of PVINs was triggered only in the most visited compartment during a 15-min exposure session. On day 3, GFP and ArchT infected mice ( $n = 6$  in both cases) were re-exposed to the place aversion apparatus during 15 min to evaluate the long-term effect of yellow-light stimulation during day 2. **b**, Time spent in the most and less visited compartments on day 1 for individual infected mice (GFP and ArchT). **c**, Average percentage of time spent in the most visited compartment on days 1,

2 and 3 for GFP- and ArchT-infected mice. A one-way analysis of variance (ANOVA) repeated measures performed on values from the GFP or the ArchT group revealed a significant effect only for the ArchT group (ArchT,  $F_{2,10} = 4.234$ ,  $P < 0.05$ ; GFP,  $F_{2,10} = 0.950$ ,  $P = 0.4191$ ). Post-hoc analysis revealed that on day 2, light inhibition of PVINs induced an aversion of the most visited compartment for ArchT infected animals in comparison to day 1 (ArchT mice, day 1 versus day 2, paired  $t$ -test,  $**P < 0.01$ ) and to GFP controls on day 2 (day 2, ArchT versus GFP, unpaired  $t$ -test,  $*P < 0.05$ ; 250-ms pulses delivered at 0.9 Hz). On day 3, ArchT mice did not avoid the most visited compartment any more (ArchT mice, day 2 versus day 3, unpaired  $t$ -test,  $*P < 0.05$ ). Error bars indicate mean  $\pm$  s.e.m.

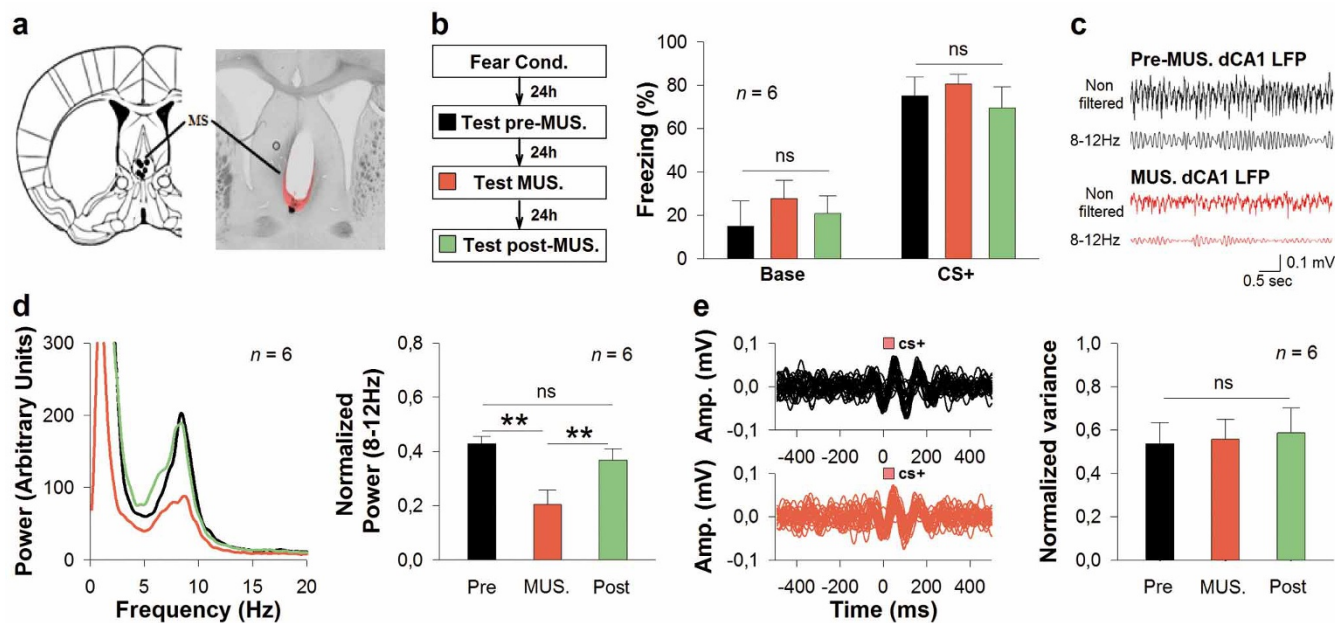


**Extended Data Figure 6 | Optogenetic activation of PVINs inhibits principal neurons and reduces freezing behaviour.** **a**, Raster plots and peristimulus time histograms illustrating the CS<sup>+</sup>-evoked excitation of a representative PN (left panel, Post FC, CS<sup>+</sup> presentations 1–4, 108 pips) and its blockade upon optogenetic-induced activation of PVINs (right panel, CS<sup>+</sup> presentations 5–8; light-pulse duration, 250 ms; 108 pips + stimulation trials) during the Post-FC session. **b**, z-score-transformed peristimulus time histogram showing PNs inhibition ( $n = 7$ ) following optogenetic-evoked activation of PVINs during CS<sup>+</sup> presentations (Post-FC session, CS<sup>+</sup> presentations 5–8; light-pulse duration, 250 ms; 108 stimulation trials). **c**, Freezing behaviour (bars,  $n = 3$  mice, block of 4 CS<sup>+</sup>) and CS<sup>+</sup>-evoked firing changes of PNs (red dots,  $n = 7$  neurons, mean z-score 100 ms post CS) before and in response to light-induced activation of PVINs during Post-FC sessions (light pulse duration, 250 ms; 108 stimulation trials; CS<sup>+</sup> 1–4 and 5–8, respectively). Optogenetic activation of PVINs inhibited PNs and reduced conditioned freezing behaviour (Wilcoxon signed-rank test,  $*P < 0.05$ ). **d**, z-score transformed peristimulus time histogram showing CS<sup>+</sup>-evoked excitation of PNs ( $n = 3$ ) exhibiting antidromic responses to BLA stimulations (Post-FC, CS<sup>-</sup> and CS<sup>+</sup> presentations 1–4, 108 pips each). These three neurons were included in the seven neurons for which CS<sup>+</sup>-evoked excitation was blocked by light excitation of PVINs (a and b). Error bars indicate mean  $\pm$  s.e.m.



**Extended Data Figure 7 | Transient amplitude increase and phase reset of local theta oscillations during fear expression.** **a**, Left panel, power spectrum of the non-filtered dmPFC LFPs recorded during Post-FC sessions ( $n = 28$  mice) for non-freezing (No freez.) and freezing (Freez.) periods showing a prominent 8–12-Hz component (that is, theta) only during non-freezing periods. Right panel, normalized theta power (8–12 Hz) for freezing and non-freezing periods during Post-FC sessions ( $n = 28$  mice, Wilcoxon signed-rank test  $***P < 0.001$ ). **b**, Top panels, non-filtered dmPFC LFP traces selected on the basis of prominent theta oscillations illustrating the transient increase in amplitude and phase reset of theta oscillations in response to CS<sup>+</sup> (Post-FC, 1 trial). Bottom-left panel, representative dmPFC 8–12-Hz LFP traces illustrating the phase reset and transient amplitude increase of theta oscillations in response to CS<sup>+</sup> or CS<sup>-</sup> presentations (Post-FC, 27 pips each). Bottom-right panel, average ratio of LFP theta power (500 ms post CS or 500 ms pre CS) in response to CS<sup>-</sup> and CS<sup>+</sup> pips. This analysis revealed a larger transient

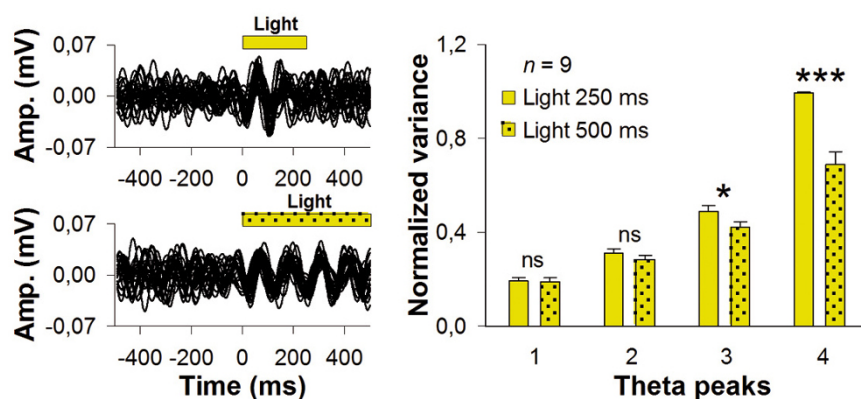
increase in LFP upon CS<sup>+</sup> presentations (Post-FC,  $n = 28$  mice, CS<sup>-</sup> versus CS<sup>+</sup>, paired  $t$ -test,  $***P < 0.001$ ). **c**, Left panel, representative dmPFC LFP traces filtered in the 8–12-Hz range, illustrating the phase resetting of theta oscillations during presentations of CS pips associated with no freezing or freezing behaviour (Post-FC, 27 pips). Right panel, quantification of the variance of the first theta peak occurrence following pip presentations in freezing and non-freezing periods (Post-FC,  $n = 28$  mice, No freez. versus Freez., paired  $t$ -test,  $***P < 0.001$ ). A small variance corresponds to a strong theta phase resetting. **d**, Quantification of the time variance of the first theta peak following CS<sup>-</sup> and CS<sup>+</sup> presentations or No freez. and Freez. periods for extinction and retrieval sessions (extinction, CS<sup>-</sup> presentations and CS<sup>+</sup> presentations 1–4,  $n = 28$  mice; retrieval, CS<sup>-</sup> and CS<sup>+</sup>,  $n = 21$  mice; CS<sup>-</sup> versus CS<sup>+</sup>, paired  $t$ -test,  $***P < 0.001$ ; No freez. versus Freez., paired  $t$ -test,  $***P < 0.001$ ). Error bars indicate mean  $\pm$  s.e.m.



**Extended Data Figure 8 | Targeted reversible inactivation of the medial septum does not block conditioned fear expression or theta phase resetting.**

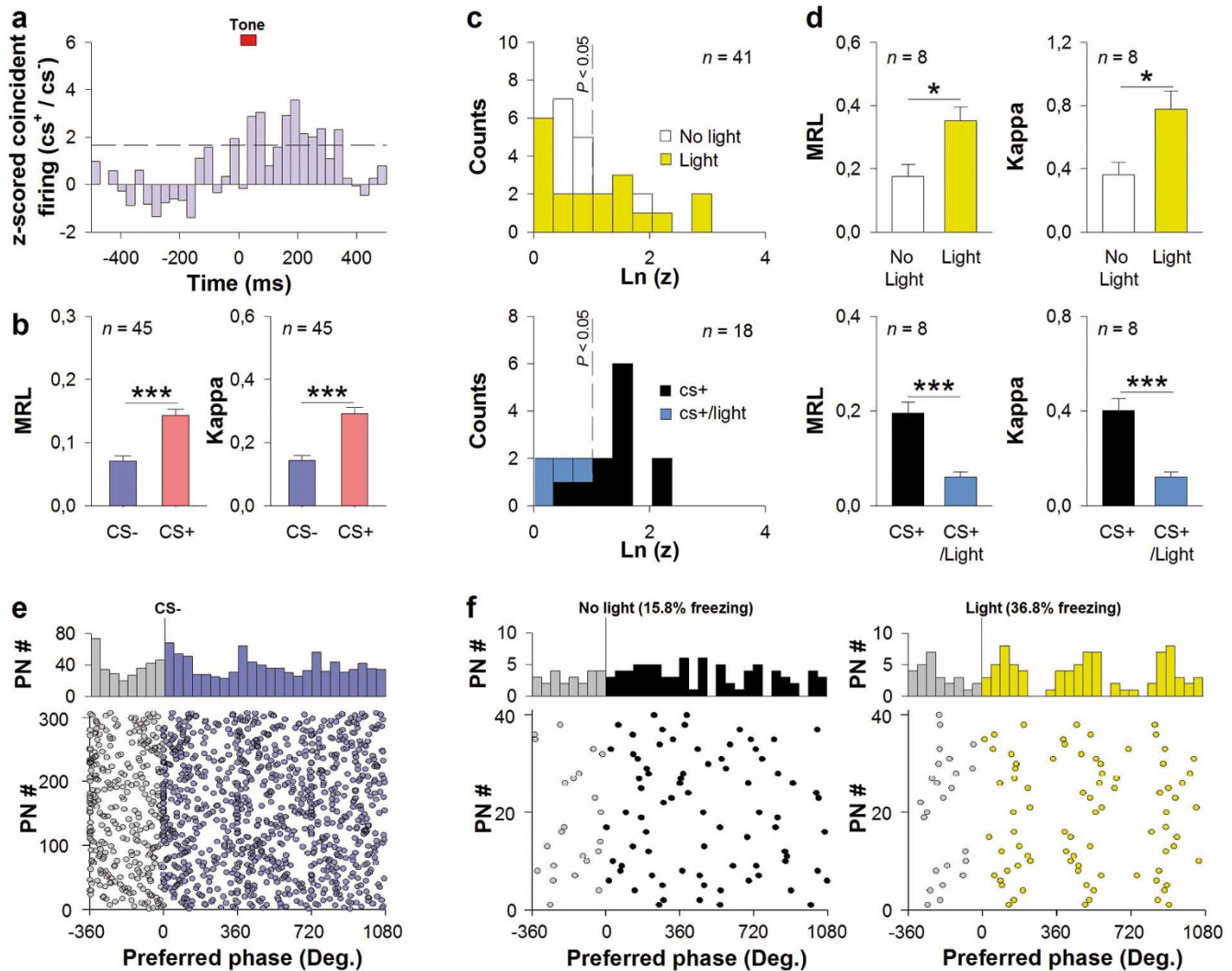
**a**, Locations of injection sites in the medial septum (MS) and composite transmission light and epifluorescence micrograph showing the MS targeted injection of muscimol (MUS.) covalently bound to a fluorescent tag (right, dipyrromethene boron difluoride (BODIPY), red). **b**, Experimental design and mean freezing values of fear conditioned mice ( $n = 6$ ) before (Test Pre-MUS.), following (Test MUS.), and one day after (Post-MUS.) injections of MUS in the MS. Following fear conditioning, targeted inactivation of the MS had no effect on basal locomotor activity or CS<sup>+</sup>-evoked freezing responses (paired  $t$ -tests). **c**, Illustrative raw and filtered (8–120-Hz) LFP traces recorded in the dorsal CA1 (dCA1) before and following MUS injections in the MS. **d**, Left panel,

power spectra of dCA1 LFPs before, following and 1 day after MS inactivation. Right panel, quantification of dCA1 LFP power (8–12 Hz) before, during and after MS inactivation. MS inactivation significantly reduced dCA1 theta power ( $n = 6$  mice, Pre-MUS. versus MUS., paired  $t$ -test,  $**P < 0.01$ ; Post-MUS. versus MUS., paired  $t$ -test,  $**P < 0.01$ , ns, not significant). **e**, Resetting of prefrontal theta oscillations. Left panel, representative dmPFC LFP traces filtered in the 8–12-Hz range (Test-MUS., first CS<sup>+</sup>). Right panel, quantification of the time variance of the first theta peak following CS<sup>+</sup> presentations before, following and 1 day after MS inactivation (Pre-MUS., MUS., Post-MUS., CS<sup>+</sup> presentations 1–4, paired  $t$ -tests). MS inactivation had no effect on dmPFC theta phase resetting upon CS<sup>+</sup> presentations. Error bars indicate mean  $\pm$  s.e.m.



**Extended Data Figure 9 | Optogenetic inhibition of prefrontal PVINs resets local theta phase.** Left panel, representative 8–12-Hz filtered LFP traces, showing the resetting of local theta phase upon optogenetic inhibition of PVINs (top part, light-pulse duration, 250 ms, 27 stimulation trials; bottom part, light-pulse duration, 500 ms, 27 stimulation trials). Right panel, quantification

of the time variance of theta peaks (theta peaks 1–4) following presentations of 250-ms or 500-ms light pulses ( $n = 9$  mice, paired  $t$ -tests,  $*P < 0.05$ ,  $***P < 0.001$ , NS, not significant). Interestingly, dmPFC theta oscillations were precisely entrained for as long as PVINs were silenced. This suggests that inhibition from PVINs masks an oscillatory process in PNs.



**Extended Data Figure 10 | PVINs control principal-neuron theta phase locking and spiking synchronization.** **a**, Normalized averaged ratio of changes in coincident activity between pairs of PNs induced by CS<sup>+</sup> and CS<sup>-</sup> and corrected for changes in firing rate (Post-FC, Ext. or Ret. sessions;  $n = 975$  pairs from 308 PNs). Dashed line indicates significant  $z$  score ( $P < 0.05$ ). Bins of 30 ms. **b**, Mean vector length (MRL) and concentration of Von Mises fit ( $\kappa$ ) upon CS<sup>-</sup> or CS<sup>+</sup>, two measures of modulation strength in phase with theta oscillations (Post-FC, Ext. or Ret. Sessions). Only neurons significantly phase locked to theta and for which at least 50 spikes were recorded during CS<sup>+</sup> were included ( $n = 45$ ) (CS<sup>-</sup> versus CS<sup>+</sup>, Wilcoxon tests, \*\*\*  $P < 0.001$ ). Error bars indicate mean  $\pm$  s.e.m. CS<sup>+</sup> entrains a stronger locking of PN spikes to ongoing theta oscillations. Together with the precise timing between CS<sup>+</sup> onset (resetting) and subsequent theta cycles, this ensures robust, coincident and timed spiking of PNs. **c**, Distribution of log-transformed Rayleigh's test  $Z$  values of PN theta modulation before and upon light-induced inhibition (top,  $n = 41$  neurons) and light-induced activation (bottom,  $n = 18$  neurons) of PVINs (light-pulse duration, 250ms; 108 trials for each; yellow light, stimulation at the end of the behavioural session; blue light, stimulation during Post-FC session, CS<sup>+</sup> presentations 5–8). Dashed line indicates significant theta phase locking threshold ( $\ln(Z) = 1.1$ ,  $P = 0.05$ ). **d**, Theta modulation of PNs significantly phase locked to theta and displaying at least 15 spikes during No light and Light conditions. Modulation with local theta was measured with the MRL (top-left panel,  $n = 8$  neurons, yellow light stimulation, paired  $t$ -tests, No light versus Light, \*  $P < 0.05$ ; bottom-left panel,  $n = 8$  neurons, blue light stimulation, No

light versus Light, \*  $P < 0.05$ ) and  $\kappa$  (top-right panel,  $n = 8$  neurons, yellow light stimulation, paired  $t$ -tests, No light versus Light, \*\*\*  $P < 0.001$ ; bottom-left panel,  $n = 8$  neurons, blue light stimulation, No light versus Light, \*\*\*  $P < 0.001$ ). Error bars indicate mean  $\pm$  s.e.m. These results show that inhibiting PVINs is both sufficient to increase PNs' modulation with local theta, and necessary for theta entrainment of PNs evoked by CS<sup>+</sup>. **e**, Top panel, distribution of PNs' preferred theta phase ( $n = 308$ ) during cycles around CS<sup>-</sup>. The phases of LFPs were aligned to the first theta peak following CS<sup>-</sup> onset to mimic phase resetting of local theta (one theta cycle before, and three theta cycles following CS were included, bins of 45°). Bottom panel, distribution of individual PNs' preferred theta phases during theta cycles around CS<sup>-</sup> showing a synchronization of PNs around the peak of the LFP (Rayleigh's test for circular uniformity, first theta cycle post CS,  $P < 0.001$ , indicating that the circular distribution is not uniform). **f**, Top panel, distribution of PNs' preferred theta phase ( $n = 41$ ) during theta cycles outside light stimulation (left part, 15.8% freezing) and upon light-induced resetting of theta oscillations (right part, 36.8% freezing; one theta cycle before, and 3 theta cycles following CS were included, bins of 45°). Bottom panel, distributions of individual PNs' preferred theta phase outside and upon light stimulation. Despite a low number of neurons and a moderate freezing induced by light inhibition of dmPFC PVINs (36.8% freezing), this analysis revealed that light-induced reset of local theta oscillations promotes neuronal synchronization of PNs (Rayleigh's test for circular uniformity, first theta cycle post CS; Light,  $P < 0.001$ ; No light,  $P = \text{NS}$ ).