ARTICLE

Stochastic synaptic plasticity underlying compulsion in a model of addiction

Vincent Pascoli[1], Agnès Hiver[1], Ruud Van Zessen[1], Michaël Loureiro[1], Ridouane Achargui[1], Masaya Harada[1], Jérôme Flakowski[1] & Christian Lüscher[1,2]*

Activation of the mesolimbic dopamine system reinforces goal-directed behaviours. With repetitive stimulation—for example, by chronic drug abuse—the reinforcement may become compulsive and intake continues even in the face of major negative consequences. Here we gave mice the opportunity to optogenetically self-stimulate dopaminergic neurons and observed that only a fraction of mice persevered if they had to endure an electric shock. Compulsive lever pressing was associated with an activity peak in the projection terminals from the orbitofrontal cortex (OFC) to the dorsal striatum. Although brief inhibition of OFC neurons temporarily relieved compulsive reinforcement, we found that transmission from the OFC to the striatum was permanently potentiated in persevering mice. To establish causality, we potentiated these synapses in vivo in mice that stopped optogenetic self-stimulation of dopamine neurons because of punishment; this led to compulsive lever pressing, whereas depotentiation in persevering mice had the converse effect. In summary, synaptic potentiation of transmission from the OFC to the dorsal striatum drives compulsive reinforcement, a defining symptom of addiction.

All addictive drugs target the mesolimbic dopamine system[1]. Through distinct cellular mechanisms, these drugs increase dopamine levels[2] even in the absence of a reward-prediction error, resulting in an excessive learning signal[3]. This may lead to loss of control, such that some individuals will shift to compulsive drug intake[4,5], used by some to define addiction[6–8]. About 20% of users of addictive substances such as cocaine, heroin and amphetamines eventually fulfil this diagnostic criterion[9].

The neural correlate for compulsive reinforcement is poorly understood but an imbalance in the systems that control goal-directed and habitual actions has been implicated[5,10,11], possibly driven by an activity shift from the ventro-medial to dorso-lateral striatum[12]. The formation of stimulus–response associations may trigger motor programs in the dorso-lateral striatum that favour habitual drug use[12]. Alternatively, addicts may suffer from a failure of 'top-down' inhibition of stimulus–response associations, a function attributed to the medial prefrontal cortex[13,14]. Finally, drugs may perturb goal-directed outcome-representation ascribed to the OFC. Indeed, pharmacological inhibition of the dorsal striatum reduced cocaine-seeking behaviour under punishment paradigms[15], and optogenetic stimulation of the prelimbic cortex has the same consequence[14]. Compulsive self-administration of cocaine is associated with enhanced activity in the OFC and inhibition of the OFC reduces compulsive reinforcement[16,17]. Moreover, the function of the OFC is disrupted after withdrawal from cocaine self-administration in rats[18–20].

Although the OFC and the striatum emerge as hubs for compulsive reinforcement, the cellular substrate that maintains reward-seeking behaviours despite negative consequences remains unknown.

Because an increase in the levels of mesolimbic dopamine is the defining commonality of addictive drugs, we implemented a model in which the mouse presses a lever to optogenetically activate the ventral tegmental area (VTA) dopaminergic neurons (optogenetic dopamine-neuron self-stimulation; oDASS), which mimics drug-induced circuit-wide adaptations[16,21]. Here we identify a cellular correlate of compulsive reinforcement, by introducing a punishment once oDASS was established.

Compulsive self-stimulation of dopamine neurons

We expressed channelrhodopsin-2 (ChR2) bilaterally in dopaminergic neurons of the VTA and implanted an optic fibre that was aimed at the midbrain (oDASS mice, see Methods and Extended Data Fig. 1). When a mouse pressed the active lever, laser stimulation (30 bursts of 5 pulses of 4 ms at 20 Hz) began after a 5-s delay. Every four days, the number of lever presses required to trigger stimulation was increased to reach a final fixed ratio of three (FR3). All mice quickly reached a maximum of 80 laser stimulations in less than one hour (Fig. 1a, b).

Following acquisition, the perseverance of oDASS despite electric foot shocks was used to evaluate compulsivity. The shock was delivered every third completed fixed ratio schedule and its intensity (500 ms, 0.25 mA) was sufficient to suppress lever pressing for sucrose reward[16]. This punishment reduced the laser-stimulation rate (Fig. 1b), albeit with high variability between individual mice. Some mice almost stopped responding, whereas others kept obtaining the maximum oDASS, taking only slightly more time (Extended Data Fig. 2). For the 109 mice, the histogram for the oDASS rate was unimodal during the baseline sessions but became bimodal by the end of the fourth punished session (Fig. 1c). A clustering method on the entire behavioural dataset revealed two distinct classes (Extended Data Fig. 3): mice with a small decrease in oDASS rate during punished sessions ($n = 66$) and mice that strongly decreased oDASS ($n = 43$), which we called perseverers and renouncers, respectively. Because we found similar proportions of male and female mice in both clusters, all subsequent experiments were carried out on both sexes (Fig. 1d). When subjected to additional punishment sessions, mice remained in the original cluster (Extended Data Fig. 4a). Moreover, there was no correlation between perseverance and baseline oDASS rate or motivation for oDASS (Fig. 1d, e and Extended Data Fig. 4b, c). Thus, in about 60% of the mice, the burst activity elicited by oDASS

[1]Department of Basic Neurosciences, Faculty of Medicine, University of Geneva, Geneva, Switzerland. [2]Clinic of Neurology, Department of Clinical Neurosciences, Geneva University Hospital, Geneva, Switzerland. *e-mail: Christian.Luscher@unige.ch
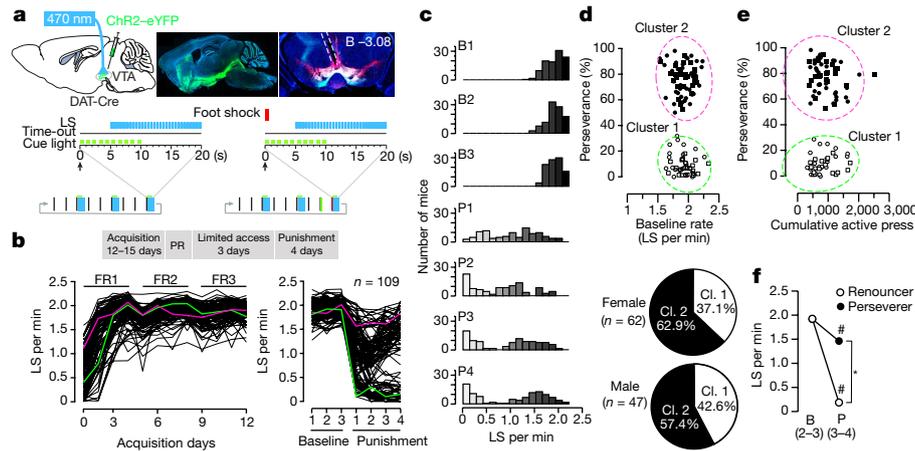
Fig. 1 | Perseverance of oDASS despite punishment. a, Schematic of optic fibre placement above the VTA of DAT-Cre mouse infected with AAV5-EF1α-DIO-ChR2-eYFP (left), image of parasagittal oblique slice and tyrosine hydroxylase staining of a midbrain coronal slice (right). B, bregma; values indicate coordinates from bregma. Mice pressed a lever during a sequence consisting of three FR3 trials (a total of 9 lever presses) to self-stimulate. Completion of a FR3 trial triggered a cue light and 5 s later the laser stimulation (LS, 30 bursts composed of 5 short laser pulses, 4-ms width at 20 Hz) as well as a time-out period during which lever presses had no consequences. During punished sessions for every third trial, the cage light was turned on for 1 s at the second press (number 8) of the FR3 and the final press (number 9) triggered a foot shock (0.25 mA, 500 ms). **b**, Mice were subjected to punishment sessions after 12 days of acquisition (maximum of 80 oDASS per day), a progressive ratio session and 3 days of limited access (maximum of 40 oDASS per day). PR, progressive ratio. oDASS rate (LS per min) during acquisition (left), baseline and punished sessions (right) for all mice ($n = 109$ mice). Green and pink lines identify the examples of two mice displayed in

Extended Data Fig. 2. **c**, Histograms of the proportion of mice binned by oDASS rate during baseline (B1–B3) and punished sessions (P1–P4) ($n = 109$ mice). **d**, Perseverance (rate of oDASS during punished sessions 3–4, normalized to baseline) as a function of the baseline rate for male and female mice, showing the two clusters identified in Extended Data Fig. 3. Pie charts show the proportion of male and female mice in clusters 1 and 2 ($n = 109$ mice). **e**, Perseverance as a function of oDASS motivation, measured by the cumulative number of active presses during a progressive ratio schedule session ($n = 98$ mice). **f**, oDASS rate (mean ± s.e.m.) during the two last sessions of baseline and punished for renouncing and persevering mice (analysis of variance (ANOVA) followed by two-sided t-test: $t_{107} = 0.11$, $P > 0.99$ and $t_{107} = 33.69$, $P < 0.0001$; *$P < 0.05$ for persevering versus renouncing mice for the punished session; $n = 66$ and 43 mice, respectively; $t_{42} = 50.24$, $P < 0.0001$ and $t_{65} = 16.26$, $P < 0.0001$; #$P < 0.05$; baseline compared to punished sessions for renouncing and persevering mice). See Supplementary Table 1 for complete statistics. **a**, Line drawing modified from Paxinos and Franklin[44], copyright © 2007.

was sufficient to stochastically induce perseverance despite negative consequences.

At baseline, the average delay between completion of FR sequences and the initiation of a subsequent trial was less than 10 s (Extended Data Fig. 4d). Once punishment was introduced, renouncers showed a strong increase in the delay to initiate the next sequence. After a completed FR trial without foot shock the delays became shorter, suggesting that continuous updating of reward value in light of the preceding outcome occurred.

The longer delays led to a reduced oDASS rate over the entire session that was different between the two clusters, dropping to less than 20% of baseline in renouncers, compared to about 80% in persevering mice (Fig. 1f).

The OFC projects to the striatum

Following a previously published screen[16], we next mapped the circuit that originates in the lateral OFC. An adeno-associated virus (AAV8-hSyn-chrimson-tdTomato) that was injected in the OFC anterogradely labelled fibre terminals in the centro-ventral part of the dorsal striatum, all along the rostro-caudal axis (Fig. 2a and Supplementary Video 1). Conversely, retrograde labelling by seeding CTB-555 in the striatum stained neurons in layers II, III and V of the OFC (Extended Data Fig. 5a).

In acute brain slices, terminal stimulation evoked large currents (400–1,000 pA) in all neurons that are located in the centro-ventral part of the striatum; these currents had both α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) and *N*-methyl-D-aspartic acid receptor (NMDAR) components (Fig. 2b). In other striatal regions, connection rates were low and current amplitudes small. When sequentially recording excitatory and inhibitory transmission, we found that the onset of inhibitory currents lagged behind excitatory currents by 2–5 ms, consistent with the existence of a feed-forward circuit[22] (Extended Data Fig. 5b). When resolving the cell-type specificity using

dopamine receptor D1R-Cre, D2R-Cre and parvalbumin-Cre mouse lines, we observed strong, converging afferents to both subtypes of spiny projection neurons (SPNs) and sparse connections onto parvalbumin interneurons (Fig. 2c, Extended Data Fig. 5c and Supplementary Video 2). Together, these experiments highlight a very strong excitatory projection from the OFC to the striatum, onto both D1R- and D2R-SPNs.

Terminal activity of OFC–striatum

We infected mice with the fluorescent calcium sensor AAV-DJ-CamKII-GCaMP6m in the OFC and placed a photometry fibre in the striatum (Fig. 3a). In baseline sessions, calcium signals decreased around lever presses in all animals (Fig. 3b and Extended Data Fig. 6a). During punished sessions, a similar decrease was observed in renouncing mice, whereas in persevering mice the calcium signal started to increase just before the lever was pressed (Fig. 3b and Extended Data Fig. 4b). In persevering mice, activity during baseline and punished sessions was therefore markedly different (Fig. 3c, d). Unpredictable foot shocks increased activity in both renouncing and persevering mice, with activity peaking after the onset of the foot shock (Extended Data Fig. 6c). Overall, the inversion of the calcium signal—which correlates with perseverance—implicates activity of the OFC–striatum pathway in compulsion.

Time-locked inhibition of OFC

We next transiently inhibited OFC neurons during oDASS to curb the increased activity observed in persevering mice. Amber light (593 nm) activated eArchT3.0 (Fig. 3e) and suppressed action potentials in slices (Fig. 3f). Inhibition in vivo immediately after the lever press that triggered the punishment-predictive cue induced a pause that eventually reduced the oDASS rate (Fig. 3g). We next inhibited OFC neurons after the completion of a sequence, which yielded a temporal profile that was similar to the profile of renouncing mice and an oDASS rate
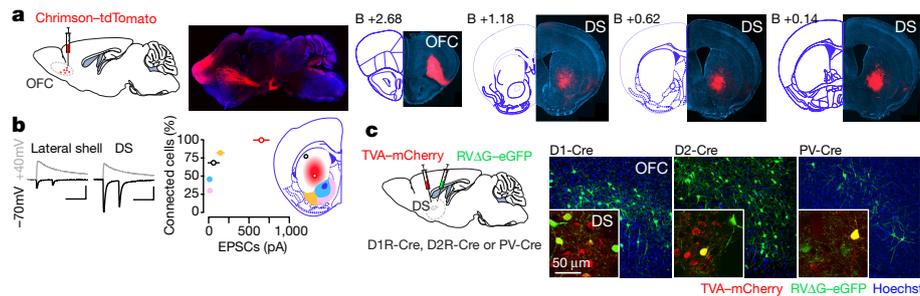
Fig. 2 | The OFC projects to the ventro-central part of the dorsal striatum. a, Schematic of the preparation for anterograde tracing from the OFC in wild-type mice. Representative sagittal and coronal images of injection site and terminals at four rostro-caudal coordinates (from bregma) of the striatum (repeated in 10 mice). DS, dorsal striatum. **b**, Left, example traces for optogenetically evoked OFC–striatum EPSCs recorded at $+40$ mV and -70 mV. Scale bars, 20 ms, 100 pA. Right, functional connectivity between OFC and striatal subregions: nucleus accumbens core (blue, $n = 13$ cells), medial and lateral shell (pink and yellow, $n = 23$ and 34 cells, respectively), dorso-lateral striatum and ventro-central striatum (black and red, $n = 16$ and 112 cells, respectively). Data are mean $\pm$ s.e.m. **c**, Schematic of retrograde tracing from specific cell types of the striatum using a rabies virus injected in transgenic mouse lines expressing Cre under the control of the dopamine D1, D2 receptor or parvalbumin (PV). Confocal pictures show retrogradely infected neurons in the OFC and starter cells in the striatum at high magnification (insets). Experiments were repeated in $n = 3$, 3 and 4 mice for D1R-Cre, D2R-Cre and PV-Cre mice, respectively. **a**, **c**, Line drawings modified from Paxinos and Franklin[44], copyright © 2007.

below 45%. The same intervention during a baseline session had no consequences for the oDASS rate (Fig. 3h). Similarly, inhibition of OFC neurons before every FR trial led to a reduction of perseverance (Extended Data Fig. 6d). In renouncing mice, inhibition of OFC neurons had little consequences for the already long delays between lever presses (Extended Data Fig. 6e).

These results suggest that OFC activity is required to invigorate compulsive oDASS, especially at the time when mice engage in the next sequence. However, the effect was transient and perseverance

returned the following day (Fig. 3i), which is why we next searched for a long-lasting alteration in the OFC output.
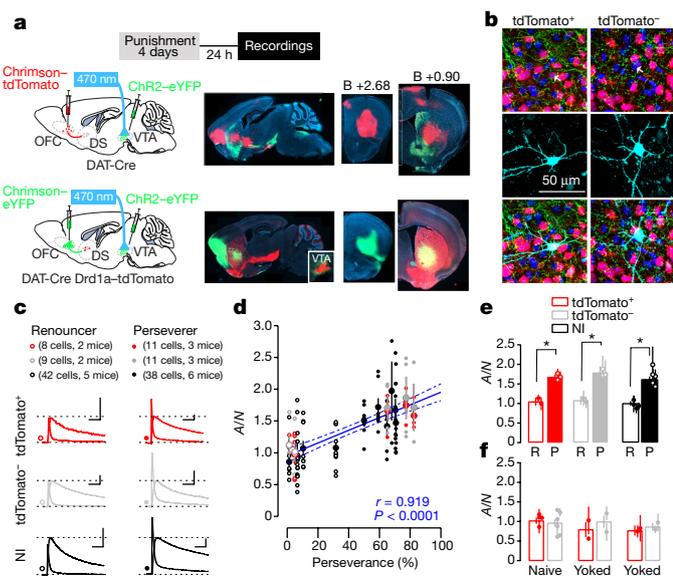
Plasticity at OFC–striatum synapses

We performed ex vivo recordings in brain slices from oDASS mice that expressed *Chrimson* in the OFC with on-the-fly identification of SPNs[23] (Fig. 4a), which was confirmed post hoc (Fig. 4b). The ratio of AMPAR to NMDAR excitatory postsynaptic current amplitudes was significantly higher in both D1R- and D2R-SPNs in persevering
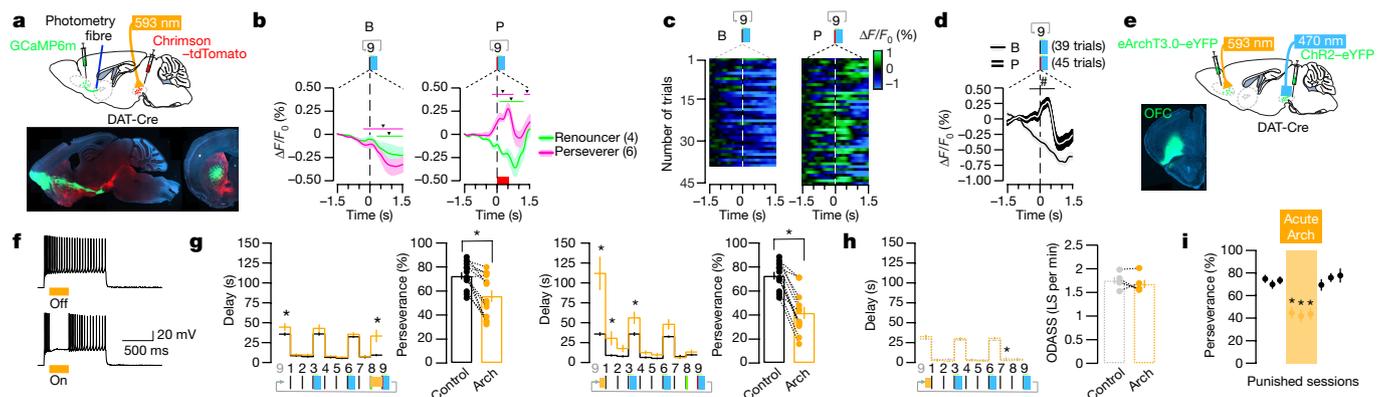
Fig. 3 | Observation and manipulation of OFC–striatum activity during oDASS. a, Schematic of the preparation of activity recordings of OFC terminals in the striatum with photometry (top) and image of a parasagittal oblique section showing GCaMP6m expression in the OFC and Chrimson–tdTomato in the VTA (repeated in 10 mice, bottom). **b**, Calcium signal ($\Delta F/F_0$) around the active lever press number 9 completing a sequence during baseline and punished sessions for renouncing and persevering mice. Red block indicates duration of electric shock. Downward triangles and green and pink lines indicate time window with significant deviation from baseline, shaded area represents s.e.m.; $n = 4$ and 6 mice. **c**, Trial activity map of calcium signals ($\Delta F/F_0$) around the active press completing the third FR3 during baseline (left) and punished (right) sessions for a persevering mouse. **d**, Group data for **c**. #Black line indicates the time window with a significant difference between punished and baseline trials using a two-sided permutation test; $n = 39$ and 45, respectively. **e**, Schematics and image of a mouse brain infected with eArchT3.0–eYFP (Arch) in the OFC and with ChR2–eYFP in the VTA (repeated in 20 mice). **f**, Activation of eArchT3.0 inhibits action potentials induced by a current step in an OFC slice. **g**, In persevering mice, OFC inhibition at specific time points during punished oDASS sessions specifically modifies the delay to engage the next action (ANOVA,

followed by two-sided t-test: *$P < 0.05$ comparing delays during punished sessions for control and eArchT3.0 stimulation). Perseverance changed as a consequence of eArchT3.0 stimulation (two-sided paired t-test: $t_{12} = 4.51$, *$P = 0.0007$, $n = 13$ mice for control and eArchT3.0 stimulation at punishment-predicted cue; $t_{11} = 8.91$, *$P < 0.0001$, $n = 12$ mice for control and eArchT3.0 stimulation after punished oDASS). **h**, Delay to engage the next action during baseline with or without OFC inhibition, with eArchT3.0 stimulation after oDASS (ANOVA followed by two-sided t-test: *$P < 0.05$ comparing delays during baseline sessions for control and eArchT3.0 stimulation). Inhibition using eArchT3.0 during baseline sessions had no consequences (two-sided paired t-test: $t_4 = 0.50$, $P = 0.64$, for oDASS rate, $n = 5$). **i**, During additional punished sessions without renewal of the intervention (inhibition after punished oDASS), the effect on perseverance was not maintained (ANOVA followed by Dunnett's test, for three consecutive comparisons: $q_{11} = 7.31$, *$P = 0.0001$; $q_{11} = 6.79$, *$P = 0.0002$; $q_{11} = 6.15$, *$P = 0.0004$; and $q_{11} = 0.96$, $P = 0.84$; $q_{11} = 0.67$, $P = 0.96$; $q_{11} = 1.23$, $P = 0.67$ for every punished versus eArchT3.0 session and punished versus recovery session, respectively; $n = 12$ mice). Data are mean $\pm$ s.e.m. for all panels. See Supplementary Table 1 for complete statistics. **a**, **e**, Line drawings modified from Paxinos and Franklin[44], copyright © 2007.

Fig. 4 | Correlation between plasticity of OFC–striatum synapses and compulsive oDASS. a, Schematic of the preparation for ex vivo recordings of OFC–striatum synapses. Recordings were performed 24 h after the fourth punishment session in slices from DAT-Cre (encoded by Dat^{cre}) mice or from $Dat^{cre}Drd1a^{tdTomato}$ mice (VTA infected with ChR2–eYFP) expressing Chrimson–tdTomato or eYFP in the OFC, respectively. Parasagittal oblique and coronal sections show infection in the OFC and terminals in the striatum (repeated in six mice). Note that VTA infection in $Dat^{cre}Drd1a^{tdTomato}$ mice is hidden by tdTomato$^+$ neurons from the striatum projecting to the midbrain. Inset shows VTA from the same animal, but from a more-medial section (repeated in three mice). **b**, In $Dat^{cre}Drd1a^{tdTomato}$ mice, recorded neurons were filled with biocytin for the identification of SPN subtypes (repeated in 39 slices from 6 mice). **c**, Average of 10 sweeps for AMPAR EPSCs in the presence of D-2-amino-5-phosphonovaleric acid (D-AP5) (50 μM) and NMDAR EPSCs isolated by subtraction for renouncing and persevering mice, with or without SPN identification with tdTomato. NI, not identified. Scale bars, 200 pA, 50 ms. **d**, AMPAR/NMDAR (A/N) ratio for every neuron as a function of perseverance, per animal and correlation (Pearson's $r = 0.919$, $P < 0.0001$; $n = 119$ cells from 16 mice). **e**, Mean AMPAR/NMDAR ratio for renouncing (R) and persevering (P) mice (ANOVA followed by two-sided t-test: $t_{17} = 3.69$, $*P = 0.002$ and $t_{18} = 4.20$ $*P = 0.0003$ for tdTomato$^+$ and tdTomato$^-$, respectively; $t_{78} = 6.72$, $*P < 0.0001$, renouncing versus persevering for unidentified neurons, same sample size as in **c**). **f**, AMPAR/NMDAR ratio for naive mice and for mice yoked to renouncing or persevering mice, with Drd1a–tdTomato identification (12 tdTomato$^+$ and 12 tdTomato$^-$ cells from, respectively, 4 and 6 naive mice; 6 tdTomato$^+$ and 9 tdTomato$^-$ cells from 2 mice yoked to renouncing mice and 7 tdTomato$^+$ and 8 tdTomato$^-$ cells from 2 mice yoked to persevering mice). Data are mean ± s.e.m. in all panels. See Supplementary Table 1 for complete statistics. **a**, Line drawings modified from Paxinos and Franklin[44], copyright © 2007.

compared to renouncing mice. In fact, there was a strong correlation between the mean AMPAR/NMDAR ratio and the perseverance (Fig. 4c–e and Extended Data Fig. 7a). In animals yoked to renouncing or persevering mice (that is, mice that only receive the shock, but have never experienced oDASS), the AMPAR/NMDAR ratio was not different from naive mice, demonstrating that the plasticity did not reflect the number of shocks received (Fig. 4f and Extended Data Fig. 7a). The release probability was also higher in persevering mice, as determined by a decrease in the paired-pulse ratio (Extended Data Fig. 7b); moreover, no change in the composition of AMPAR subunits was detected (Extended Data Fig. 6c). Because the ratio of excitatory to inhibitory postsynaptic currents was increased to the same extent as the AMPAR/NMDAR ratio in persevering mice, the OFC to interneuron synapses probably remained unaffected (Extended Data Fig. 7d). Taken together, perseverance was associated with a strengthening of OFC–striatum transmission onto SPNs.

Bidirectional shift in synaptic strength

We next tested whether potentiation of OFC–striatum transmission would lead to perseverance in renouncing mice and, conversely, whether depotentiation in persevering mice would reduce compulsivity.

We found that brief stimulation at 20 Hz of the OFC to dorsal striatum projections was sufficient to potentiate AMPAR excitatory postsynaptic currents (EPSCs) in slices from renouncing (198 ± 23%) or naive (180 ± 20%) mice, whereas the long-term potentiation (LTP) was occluded in slices from persevering mice (111 ± 15%; Fig. 5a). This protocol was then delivered in vivo through optic fibres that targeted OFC terminals in the striatum, and the efficacy was verified ex vivo by measuring an increased AMPAR/NMDAR ratio in slices (Fig. 5b and Extended Data Fig. 8a). In renouncing mice, when applied before the punishment sessions, this procedure led to a significant reduction in the delay to engage in a new sequence and strongly increased the overall oDASS rate (Fig. 5c). This behavioural change was long-lasting and observed even when punishment sessions without OFC–striatum stimulation were added (Fig. 5d). Synaptic potentiation had no effect on the baseline oDASS rate (Extended Data Fig. 8b).

To depotentiate the OFC–striatum transmission, we used two distinct approaches. First, a protocol[24] (10 Hz for 5 min) was used that restored presynaptic transmission ex vivo in slices, but that had no effect on behaviour when applied in vivo (Extended Data Fig. 9a–e). Second, low-frequency stimulation (1 Hz for 5 min, typically inducing NMDAR-dependent long-term depression (LTD) that is postsynaptically expressed[25]) in oDASS mice that expressed *Chrimson* in the OFC reliably depotentiated synapses in slices from renouncing or naive mice, while yielding very variable results in slices from persevering mice (Fig. 5e). This might be due to ambient dopamine blocking the LTD expression in D1R-SPNs[26], which was confirmed when the combination of 1-Hz stimulation and SCH23390 (a D1 receptor (D1R) antagonist) unmasked a synaptic depression in persevering mice and normalized the AMPAR/NMDAR ratio ex vivo (Fig. 5f). The paired-pulse ratio and rectification index remained unchanged (Extended Data Fig. 10a). When applied in vivo before a punished session, the protocol significantly reduced the oDASS rate in a within-control experiment, whereas separate 1-Hz simulation or SCH23390 application had no effect (Fig. 5g). oDASS was not affected when manipulations were applied before baseline sessions (Extended Data Fig. 10b). Notably, this procedure primarily affected the delay in trial initiation after punishment, similar to the temporal profile that was typically observed in renouncing mice. In contrast to the acute inhibition of OFC–striatum, the effect of LTD on behaviour was still detectable for days, even after cessation of the treatment (Fig. 5h).

Discussion

We found that in a subpopulation of mice that acquired oDASS, the strengthening of OFC–striatum synapses was causally linked to the perseverance of reinforcement despite punishment. ChR2 expression in the VTA was homogeneous and thus did not segregate with behaviour. In addition, the oDASS rate during acquisition and the breakpoint for the progressive ratio schedule were not different between mice that eventually became compulsive and those that renounced oDASS when punished. This is in contrast to previous reports in rats that self-administrated cocaine[27,28], possibly reflecting differences in the rewarding value of the drug or the training procedure. Thus, the emergence of the two behavioural phenotypes cannot be explained by differences in the intrinsic motivational properties that arise from the stimulation of dopaminergic neurons.

The transient activity of striatum-projecting OFC neurons segregated with the behavioural phenotype, which may constitute the signal for reinforced responding despite punishment. This is consistent with the OFC encoding the expected relative outcome value by updating prior information on reward and/or punishment reception[19,20,29]. Two scenarios are possible; either the value of the reward becomes excessive or the aversive nature of the punishment is discarded. On one hand, pain perception remains unaffected by oDASS[16]; on the other hand,
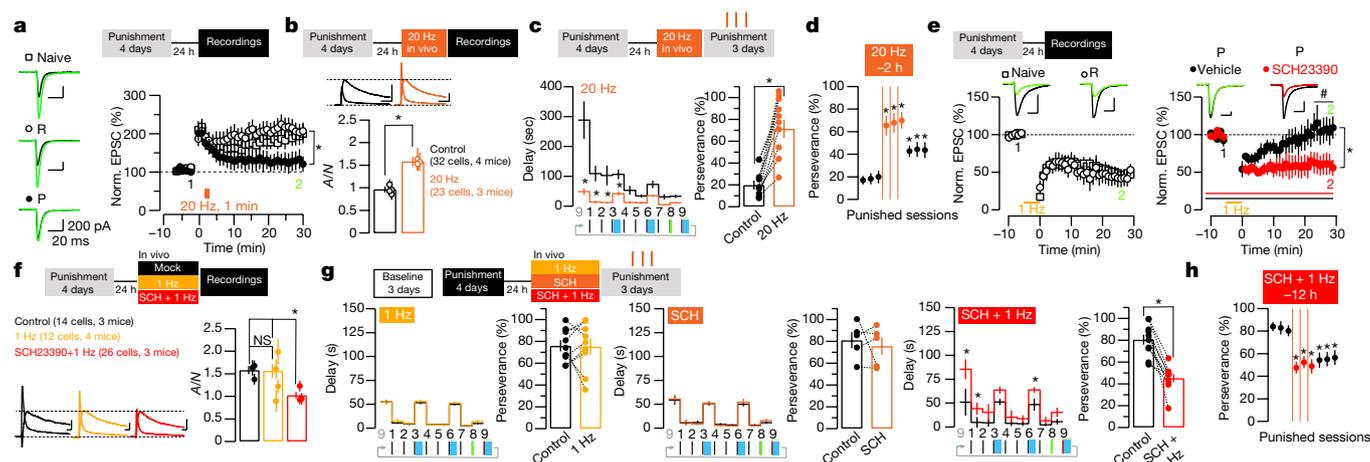
Fig. 5 | Bidirectional modulation of compulsive oDASS. a, Average traces for EPSCs recorded immediately before and 30 min after the LTP protocol (20 Hz for 1 min) and group data for normalized (norm.) EPSCs (renouncing and persevering: two-sided t-test: $t_{13} = 3.08$, $*P = 0.009$; 8 and 7 cells from 4 and 3 mice, respectively). Group data are mean ± s.e.m. of 10 cells from 3 naive mice. **b**, Ex vivo measurements of the AMPAR/NMDAR ratio after in vivo stimulation of OFC–striatum terminals at 20 Hz for 1 min in renouncing mice (control and 20 Hz: $t_{53} = 5.07$, $*P < 0.0001$; $n = 32$ cells from 4 mice and 23 cells from 3 mice, respectively). **c**, Punished sessions in renouncing mice performed 4 h after LTP induction. Delays between lever presses are reduced (ANOVA followed by two-sided t-test: $*P < 0.05$ when comparing delays during punished sessions for control and 20 Hz) and perseverance is increased (two-sided paired t-test: $t_{10} = 6.87$, $*P < 0.0001$; $n = 11$ mice for control and 20 Hz). **d**, During additional punished sessions without renewal of the intervention, perseverance remained higher (ANOVA followed by Dunnett's test: $q_{10} = 5.41$, $*P = 0.0014$; $q_{10} = 5.96$, $*P = 0.0007$; $q_{10} = 5.97$, $*P = 0.0007$; and $q_{10} = 4.31$, $*P = 0.0071$; $q_{10} = 4.19$, $*P = 0.0085$; $q_{10} = 3.64$, $*P = 0.020$; for every punished versus 20 Hz session applied 2 h before and punished versus recovery session, respectively, $n = 11$ mice). **e**, Average traces for EPSCs recorded immediately before and 30 min after the LTD protocol (1 Hz for 5 min) and group data for normalized EPSCs of naive and renouncing mice (left) and persevering mice (right). Right, in slices from persevering mice, LTD is unmasked by bath application of SCH23390 (two-sided t-test comparing persevering and renouncing mice: $t_{13} = 3.25$, $\#P = 0.006$; 104.6% for persevering mice ($n = 10$ cells

from 6 mice) versus 48.6% for renouncing mice ($n = 9$ cells from 3 mice); two-sided t-test comparing vehicle versus SCH23390: $t_{14} = 3.07$, $*P = 0.008$; 104.6% for vehicle ($n = 10$ cells from 6 mice) versus 52.6% for SCH23390 in persevering mice ($n = 8$ cells from 2 mice)). **f**, AMPAR/NMDAR ratio was normalized by in vivo stimulation of OFC–striatum terminals with 1 Hz in the presence of SCH23390 in persevering mice (ANOVA followed by two-sided t-test: $t_{24} = 0.34$, $P > 0.99$; $t_{36} = 4.41$, $*P = 0.0002$; and $t_{38} = 4.23$, $*P = 0.0003$; for control versus 1 Hz (14 and 12 cells, respectively); 1 Hz versus 1 Hz with SCH23390 (12 and 26 cells, respectively) and control versus 1 Hz with SCH23390 (14 and 26 cells, respectively)). NS, not significant. **g**, Delay between lever presses in punished sessions 12 h after SCH23390, 1 Hz or 1 Hz with SCH23390 (ANOVA followed by t-test: $*P < 0.05$ when comparing delays during punished sessions for control versus SCH23390, control versus 1 Hz or control versus1 Hz with SCH23390). Perseverance is reduced in the group of 1 Hz with SCH23390 (two-sided t-test: $t_8 = 0.1$, $P = 0.93$, $n = 9$ mice for control and 1 Hz; $t_4 = 0.6$, $P = 0.60$, $n = 5$ for control versus SCH23390; $t_9 = 6.37$, $*P = 0.0001$, $n = 10$ for control versus 1 Hz with SCH23390). **h**, During additional punished sessions without renewal of the intervention, perseverance reduction remained (ANOVA followed by Dunnett's test: $q_9 = 4.66$, $*P = 0.0054$; $q_9 = 4.43$, $*P = 0.0074$; $q_9 = 5.14$, $*P = 0.0028$; and $q_9 = 3.27$, $*P = 0.040$; $q_9 = 3.61$, $*P = 0.0024$; $q_9 = 3.72$, $*P = 0.0021$ for every punished versus 1 Hz with SCH23390 session applied 12 h before and punished versus recovery session, respectively; $n = 10$ mice). Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.

the signal revealed by fibre photometry precedes the initiation of the action and thus cannot reflect the perception of the punishment. In rats, activity of the OFC just before a lever press also correlates with compulsive cocaine self-administration[17], which may be linked to the representation of the expected reward even when there is a risk of punishment. If one accepts that the activity of the OFC increases with goal-directed actions[30] then compulsive oDASS may constitute an extreme form of goal-directed behaviour.

The main finding of our study is the identification of the synaptic strengthening of OFC–dorsal striatum projections as a neural mechanism that underlies compulsion. An important question is how does oDASS compare to cocaine exposure. Drug-evoked synaptic plasticity could be mimicked with optogenetic stimulation, starting with the strengthening of excitatory afferents onto dopaminergic neurons in the VTA after the first exposure to an addictive drug[21,31]. With chronic protocols, oDASS selectively elicits a synaptic potentiation in D1R-SPNs in the nucleus accumbens in every animal[16] that is indistinguishable from the plasticity that is observed after cocaine self-administration[32–36], both of which are associated with cue-evoked seeking behaviour. By contrast, the plasticity described here at OFC–striatum synapses was stochastically observed. Differences in the susceptibility rate to compulsive behaviour that was seen in oDASS versus cocaine self-administration (20% for cocaine self-administration[37] versus 60% found here) could also reflect different dopamine kinetics and recruited targets (for example, blocking serotonin reuptake) in the two paradigms[38].

Another open question is how the plasticity that underlies compulsion is induced. Because the innervation of the striatum by dopaminergic neurons from the VTA is weak, induction is unlikely to be a direct consequence of dopamine release. Moreover, the rules of plasticity dictated by dopamine modulation are not compatible with an induction in both D1R- and D2R-expressing SPNs[26]. Regardless, the compulsivity-associated plasticity in the striatum acts in concert with the involvement of more and more dorsal regions of the striatum as addiction develops[39].

Why only a fraction of mice lose control remains to be determined; the emergence of the two groups is even more surprising given the high degree of genetic homogeneity of the mouse line used here (our Dat^{cre} (also known as $Slc6a3^{cre}$) mice were backcrossed for more than ten generations into the C57BL/6J mouse line), and may reflect a case of stochastic individuality[40]. The emerging circuit model may help to guide molecular investigations while taking into account life experiences. For example, impulsivity has been proposed to be a predictive endophenotype for addiction[41].

The identification of adaptation of a cortical circuit that underlies the late stage of addiction enables a rational refinement of the therapeutic interventions that are currently tested in people with addiction using pharmacology, deep brain stimulation or transcranial magnetic stimulation[42,43].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0789-4

1. Lüscher, C. & Ungless, M. A. The mechanistic classification of addictive drugs. *PLoS Med.* **3**, e437 (2006).
2. Di Chiara, G. et al. Dopamine and drug addiction: the nucleus accumbens shell connection. *Neuropharmacology* **47**, 227–241 (2004).
3. Keiflin, R. & Janak, P. H. Dopamine Prediction errors in reward learning and addiction: from theory to neural circuitry. *Neuron* **88**, 247–263 (2015).
4. Koob, G. F. Antireward, compulsivity, and addiction: seminal contributions of Dr. Athina Markou to motivational dysregulation in addiction. *Psychopharmacology* **234**, 1315–1332 (2017).
5. Smith, R. J. & Laiks, L. S. Behavioral and neural mechanisms underlying habitual and compulsive drug seeking. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **87**, 11–21 (2018).
6. Vanderschuren, L. J. M. J. & Everitt, B. J. Behavioral and neural mechanisms of compulsive drug seeking. *Eur. J. Pharmacol.* **526**, 77–88 (2005).
7. Volkow, N. D., Koob, G. F. & McLellan, A. T. Neurobiologic advances from the brain disease model of addiction. *N. Engl. J. Med.* **374**, 363–371 (2016).
8. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, compulsivity, and top-down cognitive control. *Neuron* **69**, 680–694 (2011).
9. Yücel, M. et al. A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. *Addiction* https://doi.org/10.1111/add.14424 (2018).
10. Everitt, B. J. & Robbins, T. W. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* **8**, 1481–1489 (2005).
11. Vandaele, Y. & Janak, P. H. Defining the place of habit in substance use disorders. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **87**, 22–32 (2017)
12. Everitt, B. J. & Robbins, T. W. Drug addiction: updating actions to habits to compulsions ten years on. *Annu. Rev. Psychol.* **67**, 23–50 (2016).
13. McCracken, C. B. & Grace, A. A. Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization. *J. Neurosci.* **33**, 17469–17482 (2013).
14. Chen, B. T. et al. Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature* **496**, 359–362 (2013).
15. Jonkman, S., Pelloux, Y. & Everitt, B. J. Differential roles of the dorsolateral and midlateral striatum in punished cocaine seeking. *J. Neurosci.* **32**, 4645–4650 (2012).
16. Pascoli, V. et al. Sufficiency of mesolimbic dopamine neuron stimulation for the progression to addiction. *Neuron* **88**, 1054–1066 (2015).
17. Guillem, K. & Ahmed, S. H. Preference for cocaine is represented in the orbitofrontal cortex by an increased proportion of cocaine use-coding neurons. *Cereb. Cortex* **28**, 819–832 (2018).
18. Lucantonio, F., Stalnaker, T. A., Shaham, Y., Niv, Y. & Schoenbaum, G. The impact of orbitofrontal dysfunction on cocaine addiction. *Nat. Neurosci.* **15**, 358–366 (2012).
19. Lucantonio, F. et al. Effects of prior cocaine versus morphine or heroin self-administration on extinction learning driven by overexpectation versus omission of reward. *Biol. Psychiatry* **77**, 912–920 (2015).
20. Schoenbaum, G., Chang, C. Y., Lucantonio, F. & Takahashi, Y. K. Thinking outside the box: orbitofrontal cortex, imagination, and how we can treat addiction. *Neuropsychopharmacology* **41**, 2966–2976 (2016).
21. Brown, M. T. C., Korn, C. & Lüscher, C. Mimicking synaptic effects of addictive drugs with selective dopamine neuron stimulation. *Channels* **5**, 461–463 (2011).
22. Sciamanna, G., Ponterio, G., Mandolesi, G., Bonsi, P. & Pisani, A. Optogenetic stimulation reveals distinct modulatory properties of thalamostriatal vs corticostriatal glutamatergic inputs to fast-spiking interneurons. *Sci. Rep.* **5**, 16742 (2015).
23. Gerfen, C. R. et al. D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250**, 1429–1432 (1990).
24. Grueter, B. A., Brasnjo, G. & Malenka, R. C. Postsynaptic TRPV1 triggers cell type-specific long-term depression in the nucleus accumbens. *Nat. Neurosci.* **13**, 1519–1525 (2010).
25. Pascoli, V., Turiault, M. & Lüscher, C. Reversal of cocaine-evoked synaptic potentiation resets drug-induced adaptive behaviour. *Nature* **481**, 71–75 (2012).
26. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
27. Pelloux, Y., Everitt, B. J. & Dickinson, A. Compulsive drug seeking by rats under punishment: effects of drug taking history. *Psychopharmacology* **194**, 127–137 (2007).
28. Kasanetz, F. et al. Transition to addiction is associated with a persistent impairment in synaptic plasticity. *Science* **328**, 1709–1712 (2010).
29. Lucantonio, F. et al. Orbitofrontal activation restores insight lost after cocaine use. *Nat. Neurosci.* **17**, 1092–1099 (2014).
30. Padoa-Schioppa, C. & Conen, K. E. Orbitofrontal cortex: a neural circuit for economic decisions. *Neuron* **96**, 736–754 (2017).
31. Ungless, M. A., Whistler, J. L., Malenka, R. C. & Bonci, A. Single cocaine exposure *in vivo* induces long-term potentiation in dopamine neurons. *Nature* **411**, 583–587 (2001).
32. Pascoli, V. et al. Contrasting forms of cocaine-evoked plasticity control components of relapse. *Nature* **509**, 459–464 (2014).
33. Terrier, J., Lüscher, C. & Pascoli, V. Cell-type specific insertion of GluA2-lacking AMPARs with cocaine exposure leading to sensitization, cue-induced seeking, and incubation of craving. *Neuropsychopharmacology* **41**, 1779–1789 (2016).
34. Hearing, M., Graziane, N., Dong, Y. & Thomas, M. J. Opioid and psychostimulant plasticity: targeting overlap in nucleus accumbens glutamate signaling. *Trends Pharmacol. Sci.* **39**, 276–294 (2018).
35. Lüscher, C. The emergence of a circuit model for addiction. *Annu. Rev. Neurosci.* **39**, 257–276 (2016).
36. Wolf, M. E. Synaptic mechanisms underlying persistent cocaine craving. *Nat. Rev. Neurosci.* **17**, 351–365 (2016).
37. Deroche-Gamonet, V., Belin, D. & Piazza, P. V. Evidence for addiction-like behavior in the rat. *Science* **305**, 1014–1017 (2004).
38. Pelloux, Y., Dilleen, R., Economidou, D., Theobald, D. & Everitt, B. J. Reduced forebrain serotonin transmission is causally involved in the development of compulsive cocaine seeking in rats. *Neuropsychopharmacology* **37**, 2505–2514 (2012).
39. Belin, D. & Everitt, B. J. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* **57**, 432–441 (2008).
40. Honegger, K. & de Bivort, B. Stochasticity, individuality and behavior. *Curr. Biol.* **28**, R8–R12 (2018).
41. Belin, D., Mar, A. C., Dalley, J. W., Robbins, T. W. & Everitt, B. J. High impulsivity predicts the switch to compulsive cocaine-taking. *Science* **320**, 1352–1355 (2008).
42. Diana, M. et al. Rehabilitating the addicted brain with transcranial magnetic stimulation. *Nat. Rev. Neurosci.* **18**, 685–693 (2017).
43. Coles, A. S., Kozak, K. & George, T. P. A review of brain stimulation methods to treat substance use disorders. *Am. J. Addict.* **27**, 71–91 (2018).
44. Paxinos, G. & Franklin, K. B. J. *The Mouse Brain in Stereotaxic Coordinates* (Academic, New York, 2007).

Author contributions V.P. conceived the experiments and performed patch recordings and behavioural experiments. A.H. did surgeries for viral infection, behavioural experiments and in vivo recordings. R.A. carried out the retrograde tracing with the rabies strategy. R.V.Z. implanted the photometry and carried out analyses and recordings. M.L. and M.H. carried out patch recordings. J.F. carried out clustering analysis. C.L. conceptualized and supervised the study, and prepared the manuscript with the help of all authors.

Competing interests C.L. is a member of the following scientific advisory boards: Stalicla SA; Phénix Foundation; International research in paraplegia (IRP) Foundation, Geneva.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0789-4.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0789-4.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to C.L.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Animals. Mice (age 8–24 weeks) were heterozygous BAC-transgenic mice in which the Cre-recombinase expression was under the control of the regulatory elements of the dopamine transporter gene (DAT-Cre mice[45]). Dat^{cre} mice were originally provided by G. Schutz and only heterozygous mice were used for experiments. DAT B6.SJL-Slc6a3tm1.1(cre)Bkmn/J (also known as DAT-IRES-Cre) mice crossed with mice in which tomato expression was driven by the D1R ($Drd1a^{tdTomato}$ from Jackson Laboratories) gene regulatory element were also used. Tg(Drd1-cre)120Mxu, Tg(Drd2-cre)ER43Gsat and 129P2-Pvalbtm1(cre)Arbr/J mice from Charles River were used for tracing studies. Weights and genders were distributed homogeneously among the groups. Transgenic mice had been backcrossed into the C57BL/6 line for a minimum of four generations. Mice were single-housed after surgery. All animals were kept in a temperature- and humidity-controlled environment with a 12-h light/12-h dark cycle (lights on at 07:00). All procedures were approved by the Institutional Animal Care and Use Committee of the University of Geneva and by the animal welfare committee of the Cantonal of Geneva, in accordance with Swiss law.

Stereotaxic injections. AAV5-EF1a-DIO-ChR2(H134R)-eYFP or AAV8-hSyn-Flex-ChrimsonR-tdTomato produced at the University of North Carolina (UNC Vector Core Facility) were injected into the VTA of 5- to 6-week-old mice. Anaesthesia was induced at 5% and maintained at 2.5% isoflurane (w/v) (Baxter) during surgery. The mouse was placed in a stereotaxic frame (Angle One) and craniotomies were performed using stereotaxic coordinates (for VTA: anterior–posterior (AP) −3.3; medial–lateral (ML) −0.9 with a 10° angle; dorsal–ventral (DV) −4.3). Injections of virus (0.5 µl) used graduated pipettes (Drummond Scientific), broken back to a tip diameter of 10–15 mm, at an infusion rate of 0.05 µl min^{-1}. Following the same procedure, AAV8-hSyn-ChrimsonR-tdTomato (UNC), AAV8-hSyn-Chrimson-GFP (Duke University), AAV5-CamKII-eArchT3.0-eYFP (UNC) or AAV-DJ-CamKII-GCaMP6m (Stanford University) was injected bilaterally in the OFC (AP +2.6; ML ±1.60; DV −1.8). During the same surgical procedure, a unique chronically indwelling optic fibre cannula was implanted above the VTA using the exact same coordinates as for the injection except for DV coordinate, which was reduced to 4.2. Three screws were fixed into the skull to support the implant, which was further secured with dental cement. The first behavioural session typically occurred 10−14 days after surgery to allow sufficient expression of ChR2 or Chrimson. A photometry fibre was implanted unilaterally in the striatum (AP +0.8; ML 1.75; DV −2.4), fibres for eArchT3.0 stimulation were placed in each OFC (AP +2.6; ML ±1.60; DV −1.6) and fibres for terminals stimulation in the striatum were implanted bilaterally (AP +0.8; ML ±1.75; DV −2.4). In D1R-Cre, D2R-Cre and PV-Cre mice, AAV5-EF1a-Flex-TVA-mCherry, AAV8-CA-Flex-RG and EnvA G-deleted Rabies-eGFP from Salk Institute were injected in the striatum (AP +0.6; ML +1.8; DV −3.2) a week before the pseudotyped rabies virus.

Optogenetic self-stimulation apparatus. Mice infected with ChR2 (or Chrimson) in the VTA were placed during the light phase in operant chambers (ENV-307A-CT, Med Associates) situated in a sound-attenuating box (Med Associates). The optic fibre of the mouse was connected to DPSS blue- or orange-light lasers (CNI-473-140-10-LED-TTL1-MM200FC; CNI-593-200-10-LED-TTL1-MM200FC; Laser 2000) via a FC/PC fibre cable (M72L02; Thorlabs) and a simple rotary joint (FRJ-FC-FC; Doric Lenses) allowing free movement during operant behaviour. Power at the exit of the patch cord was set to 15 ± 1 mW. Two retractable levers were present on one wall of the chamber and a cue light was located above each lever. A cage light was present in each chamber. A rack mount interface cabinet (SG-6010A, Med Associates) containing a programmable constant current shocker (ENV-413, Med Associates) connected to a quick disconnect grid harness (ENV-307A-QD, Med Associates) was used to provide foot shocks during punishment sessions. The apparatus was controlled and data captured using MED-PC IV software (Med Associates). For acute stimulation in the OFC or at terminals of OFC fibres in the striatum, a double rotary joint (FRJ_1x2i_FC-2FC, Doric Lenses) was used to connect cables to each hemisphere. For orange laser stimulation of eArchT3.0 or Chrimson, a shutter (CMSA-SR475_FC, Doric Lenses) was used to avoid variation in intensities during laser warm-up.

Self-stimulation acquisition and progressive ratio. Dat^{cre} mice learned to self-stimulate dopaminergic neurons in the VTA infected with AAV5-DIO-ChR2-eYFP (oDASS) for 12 consecutive days. Each of the 12 acquisition sessions lasted 120 min or until the mouse reached 80 optogenetic stimulations, whichever came first. In each oDASS session, mice could respond on the active lever, resulting in VTA stimulation. Responding on the inactive lever had no consequences. During the first 4 sessions, a single press on the active lever (termed fixed-ratio one or FR1) resulted in a 10-s illumination of a cue light (pulses of 1 s at 1 Hz). After a delay of 5 s, onset of a 15-s laser stimulation (473 nm) composed of 30 bursts separated by 250 ms (each burst consisted of 5 laser pulses of 4-ms pulse width at 20 Hz). A 20-s time-out followed the rewarded lever press, during which lever presses had no consequences but were still recorded. Next, a FR2 (sessions 5–8) and a FR3 (sessions 9–12) were introduced.

For measurements of motivation, mice did a single progressive ratio session between acquisition sessions 11 and 12 that lasted for a maximum of 4 h. The breakpoint was considered to be the last reached reinforced schedule after 4 h or after 40 min had elapsed since the last reinforced schedule. The reinforced schedules were the following: 1, 3, 5, 8, 12, 16, 22, 29, 38, 50, 65, 84, 108, 139, 178, 228, 291, 371, 473, 603, 767, 977, 1,243 and 1,582. The total number of active lever presses instead of the breakpoint was represented in Fig. 1 to better visualize individual performance.

Punishment sessions. After acquisition, mice underwent 3 additional sessions with a reduced cut-off (maximum 40 laser stimulations or 60 min, whichever came first). These sessions served as a baseline before starting the punishment sessions. Punishment sessions occurred in the exactly same conditions as for baseline sessions, except every third laser stimulation event was preceded by a foot shock (500 ms, 0.25 mA) starting immediately after the completion of the FR3 (5 s before the onset of the laser stimulation). Impending punishment was announced by illumination of the home cage light for 2-s after the second press of the FR3. Indeed, a new cue (cage light) predicting the oncoming shock was paired with the second lever press of the FR3 schedule, directly preceding the shock-coupled press.

Fibre photometry recordings. Fibre photometry recordings were performed during baseline and punished sessions in oDASS mice. A batch of DAT-Cre animals infected with Chrimson (AAV8-hSyn-Flex-ChrimsonR-tdTomato produced at UNC) in the VTA and with GCamp6m (AAV-DJ-CamKII-GCaMP6m, Stanford University) in the OFC were used for recordings of the activity of terminals in the striatum during oDASS. Mice were recorded for 20–40 min per session to minimize bleaching. OFC–striatum terminals were illuminated with blue (470 nm wavelength, M470F3, Thorlabs) and violet (405 nm wavelength, M405FP1, Thorlabs) filtered excitation LED lights, that were sinusoidally modulated at 211 and 531 Hz. Green emission light (500–550 nm) was collected through the same fibre that was used for excitation and passed onto a photoreceiver (Newport 2151, Doric Lenses). Pre-amplified signals were then demodulated by a real-time signal processor (RZ5P, Tucker Davis Systems) to determine contributions from 470 nm and 405 nm excitation sources[46]. TTL signals of the relevant stimuli were directly sent from the operant chamber to the signal processor. Analysis was performed offline in MATLAB. To calculate $\Delta F/F_0$, a linear fit was applied to the 405-nm control signal to align it to the 470-nm signal. This fitted 405-nm signal was used as F_0 in standard $\Delta F/F_0$ normalization (($F(t) - F_0(t))/F_0(t)$). Averaged peristimulus activity traces were then constructed for which the mean baseline fluorescence (−1.5 to −0.5 s before the relevant event) was subtracted from the trace.

Fibre photometry data. To identify active lever-press (ALP)-related $\Delta F/F_0$ signal modulations over time, we examined the signal across mice (renouncing versus persevering) aligned to the ALP, using a time bin of 10 ms, following a published protocol[47,48]. Baseline intervals was taken from −1.5 to −0.5 s and the event from −0.5 to 1.5 s, followed by the calculation of the average of the mean ± s.d. $\Delta F/F_0$ values of the baseline. We then checked bin-by-bin (100 ms) for a threshold of 1.65 s.d. (95% confidence interval) and significant modulation was found if 20 consecutive bins passed the threshold test.

To compare trials from the same animal, we used a permutation test[49,50]. In brief, we extracted all trials aligned on the ALP in the interval –1.5 to 1.5 s for the two conditions. We then collected all trials of the baseline and punishment sessions and randomly drew from this combined set as many trials as in the baseline session (subset 1) while the remaining trials were placed in subset 2. This random partition was repeated 1,000 times to compute the $\Delta F/F_0$ mean and s.d. for these two permutation subsets across the trials at every time point (bin size = 10 ms). Significance was achieved if 20 consecutive bins met the threshold of 3.29 s.d. (99% confidence interval).

Acute inhibition of the OFC during oDASS. Acute inhibition of the OFC transmission during oDASS sessions was performed in mice infected with AVV5-CamKII-eArchT3.0-eYFP in the OFC and ontogenetic fibres were bilaterally implanted in the OFC. Acute inhibition started at a specific epoch of behaviour and lasted for a maximum of 90 s in animals that were infected in the OFC with eArchT3.0 in the OFC. Amber laser light started at different epochs of behaviour: (1) immediately after the active lever press that triggered the punishment-predictive cue (home cage light) until the next press or for a maximum of 90 s, (2) after oDASS of the punished trial until the next press or for a maximum of 90 s, and (3) immediately after every oDASS until the next press or for a maximum of 90 s. Stimulation consisted of continuous laser activation for 6 s followed by a 2-s time-out period, repeated until the mouse initiated the next epoch or for a maximum of 90 s. Acute inhibition before the initiation of a next trial was also tested during baseline sessions.

Slice electrophysiology. Whole-cell patch-clamp recordings of striatal neurons were performed 24 h after the last punished session, after acquisition or in slices from naive mice. The AMPAR/NMDAR ratio was calculated at +40 mV with the AMPAR component pharmacologically isolated using D-AP5 (50 µM) and the NMDAR EPSC component was determined by subtraction. Currents were evoked

with optogenetic stimulation on slices of OFC terminals infected with Chrimson. When using BAC transgenic (Drd1a–tdTomato) mice crossed with Dat^{cre} mice, the OFC was infected with AAV8-hSyn-Chrimson-eYFP and tdTomato$^+$ or tdTomato$^-$ cells were filled with biocytin and identified post hoc on confocal images.

Coronal 230-μm slices of mouse brain were prepared in cooled artificial cerebrospinal fluid containing (in mM): NaCl 119, KCl 2.5, MgCl2 1.3, CaCl2 2.5, Na2HPO4 1.0, NaHCO3 26.2 and glucose 11, bubbled with 95% O2 and 5% CO2. Slices were kept at 32–34 °C in a recording chamber superfused with 2.5 ml min^{-1} artificial cerebrospinal fluid.

Ex vivo synaptic properties of the striatum. Visualized whole-cell patch-clamp recording techniques were used to measure synaptic responses to optogenetic stimulation of OFC terminals. In some experiments, D1R- and D2R-SPNs of the striatum were identified by the presence of the tdTomato in BAC transgenic mice by using a fluorescence microscope (Olympus BX50WI, fluorescent light U-RFL-T) and confirmed on confocal images of the recorded neuron filled with biocytin (Sigma, B4261) and stained with streptavidin–Cy5 (Invitrogen, 434316). The holding potential was −70 mV and the access resistance was monitored by a hyperpolarizing step of −14 mV. The liquid junction potential was small (−3 mV), and therefore traces were not corrected. Experiments were discarded if the access resistance varied by more than 20%. Currents were amplified (Multiclamp 700B, Axon Instruments), filtered at 5 kHz and digitized at 20 kHz (National Instruments Board PCI-MIO-16E4, Igor, Wave Metrics). For recordings of optogenetically evoked EPSCs, the internal solution contained (in mM): CsCl 130, NaCl 4, creatine phosphate 5, MgCl2 2, NA2ATP 2, NA3GTP 0.6, EGTA 1.1, HEPES 5 and spermine 0.1. QX-314 (5 mM) was added to the solution to prevent action currents. Synaptic currents were evoked by short light pulses (4 ms) at 0.1 Hz through an LED (M590L3-C1, ThorLabs) placed through the objective above the tissue. To isolate AMPAR-evoked EPSCs the NMDA antagonist D-2-amino-5-phosphonovaleric acid (D-AP5, 50 μM) was applied to the bath. The NMDAR component was calculated as the difference between the EPSCs measured in the absence and in the presence of D-AP5. The AMPAR/NMDAR ratio was calculated by dividing the peak amplitudes. The AMPAR/NMDAR ratio was also calculated by taking the NMDAR EPSC component 20 ms after the peak of the EPSCs recorded at +40 mV without pharmacological isolation. The rectification index of AMPAR was calculated as the ratio of the chord conductance calculated at negative potential divided by chord conductance at positive potential. The paired-pulse ratio (PPR) was measured during the first 3 min of the recordings by delivering 2 pulses of 4 ms with a 76-ms interval. Examples traces are averages of 10–15 sweeps. All experiments were performed in the presence of picrotoxin (100 μM).

In vitro synaptic plasticity. Low-frequency stimulation (1 or 10 Hz for 5 min) was applied with 4-ms light pulses and the magnitude of LTD was determined by comparing average EPSCs that were recorded 20–30 min after induction to EPSCs recorded immediately before induction. These experiments were conducted with bath application of the corresponding vehicle, with the D1R-antagonist SCH23390 (10 μM) or with the NMDA use-dependent channel blocker MK801 (10 μM) and the mGluR5-positive allosteric modulator (PAM, CBPPB 30 μM). For LTD on slices, the internal solution contained (in mM) CsCl 130, NaCl 4, creatine phosphate 5, MgCl2 2, NA2ATP 2, NA3GTP 0.6, EGTA 1.1, HEPES 5, QX-314 5 and spermine 0.1. For induction of a LTP Chrimson expressed in OFC terminals was stimulated at 20 Hz for 1 min. For LTP experiments, the following internal solution was used (in mM): potassium gluconate 140, MgCl2 2, KCl 5, Na2ATP 4, Na3GTP 0.3, creatine phosphate 10, HEPES 10 and EGTA 0.2. For validation of terminal inhibition with 3 Hz stimulation for 1 min the internal solution was (in mM): CsCl 130, NaCl 4, creatine phosphate 5, MgCl2 2, NA2ATP 2, NA3GTP 0.6, EGTA 1.1, HEPES 5, QX-314 5 and spermine 0.1. EPSCs were evoked at 0.1 Hz before and after the protocol (3-Hz stimulation). All experiments were performed in the presence of picrotoxin (100 μM).

Feed-forward inhibition. For recordings of transmission, EPSCs and inhibitory postsynaptic currents (IPSCs) from OFC to SPNs of the striatum, the internal solution contained (in mM): CsCH3SO4 128, NaCl 20, CaCl2 0.3, MgCl2 1, NA2ATP 2, NA3GTP 0.3, EGTA 1 and HEPES 10. The holding potential was −70 mV for EPSC recordings (reversal potential for GABA (γ-aminobutyric acid)) and 0 mV for IPSC recordings (reversal potential for AMPA). For pharmacological validation, picrotoxin (100 μM) or NBQX (20 μM) was added to the bath perfusion during the recordings. The ratio of excitatory to inhibitory transmission was measured as the ratio of the charge transfer obtained at −70 mV and 0 mV for 5 light pulses and was determined at different frequencies (5, 10, 20 and 40 Hz). Charge transfer is determined as the sum of area under the curves for EPSCs or IPSCs.

eArchT3.0 validation. For recordings of OFC pyramidal neurons infected with eArchT3.0 the internal solution contained (in mM): potassium gluconate 130, MgCl2 4, Na2ATP 3.4, Na3GTP 0.1, creatine phosphate 10, HEPES 5 and EGTA 1.1. Firing was triggered by a current step (1-s duration, 200-pA step) with or without stimulation of eArchT3.0 with the amber LED. No clamp was imposed

and the cell was discarded if the resting membrane potential varied by more than 10%.

In vivo plasticity with optogenetic stimulation protocols and pharmacology. Optogenetic protocols were applied once in vivo, through bilaterally implanted optical fibres targeting the striatum 2−24 h before a punished session, or before animals were killed for ex vivo recordings. DPSS orange-light lasers (CNI-593-200-10-LED-TTL1-MM200FC, Laser 2000) connected to the indwelling optic fibre via customized patch cords (M72L02, Thorlabs) and a double rotary joint (FRJ_1x2i_FC-2FC, Doric Lenses) allowed mice to move freely during stimulation. The laser was triggered to deliver 4-ms pulses at 1 Hz or 20 Hz for 5 or 1 min, respectively. Optogenetic stimulation was applied in the home cage 4−24 h before the punished session or before the animals were killed for ex vivo electrophysiology recordings. Protocols were also tested before non-punished sessions. SCH23390 (0.3 mg kg^{-1}, 0.1% DMSO; Tocris, 0925) was given intraperitoneally (10 ml kg^{-1}), 20 min before the optogenetic stimulation protocol. CDPPB (30 mg kg^{-1}, 10% Tween 80 (Tocris, 3235) and (+)-MK801 (0.3 mg kg^{-1} (Tocris, 0924)) was given intraperitoneally (10 ml kg^{-1}), 50 and 20 min before optogenetic stimulation at 10 Hz. Different batches of mice received either the pharmacology or the optogenetic stimulation protocols before testing for perseverance in punished sessions.

Tissue preparation for imaging. Mice were anaesthetized with pentobarbital (300 mg kg^{-1}, intraperitoneally, Sanofi-Aventis) and transcardially perfused with 4% (w/v) paraformaldehyde in PBS (pH 7.5). Brains were post-fixed overnight in the same solution and stored at 4 °C. Coronal or parasagittal oblique sections (70-mm thick) were cut with a vibratome (Leica), stained with Hoechst (Sigma-Aldrich) and mounted with Mowiol (Sigma-Aldrich). Full images of brain slices were obtained with a Zeiss Axioscan Z1 system equipped with a Plan-Apochromat 10×/0.45 NA objective, together with filters for 4′,6-diamidino-2-phenylindole (DAPI) (emission band-pass filter: 445/50 nm), enhanced green fluorescent protein (eGFP) (emission band-pass filter: 525/50 nm), cyanine 3 (Cy3) (emission band-pass filter: 605/70 nm) and cyanine 5 (Cy5) (emission band-pass filter: 690/50 nm). Images from VTA, OFC and striatum were obtained using sequential laser scanning confocal microscopy (Zeiss LSM700). Photomicrographs were obtained with the following band-pass and long-pass filter settings: UV excitation (band-pass filter: 365/12 nm), GFP (band-pass filter: 450−490 nm), Cy3 (band-pass filter: 546/12 nm) and Cy5 (band-pass filter: 546/12 nm). For biocytin (Sigma-Aldrich, B4261) staining, streptavidin–Cy5 (Invitrogen 434316) was used, high-magnification images were obtained and a z-stack was made. For immunohistochemistry, the following primary antibody (rabbit polyclonal anti-tyrosine hydroxylase, Millipore AB152, lot 2722866, diluted 1:500) and the secondary antibody (donkey anti-rabbit Cy3, Millipore AP182C, lot 2397069, diluted 1:500) were used.

Clarity. C57BL/6J mice infected with AAV8-Syn-Flex-ChrimsonR-tdt in the OFC or D1R-Cre mice infected with AAV8-hSyn-Flex-TVA-P2A-GFP, AAV8-CA-Flex-RG and EnvA G-deleted Rabies-mCherry-RbE (from Salk Institute) in the striatum were anaesthetized with pentobarbital (300 mg kg^{-1}, intraperitoneally, Sanofi-Aventis) and transcardially perfused with 4% (w/v) paraformaldehyde in PBS (pH 7.5).

Brains were extracted, immersed for 12 h in 4% PFA, rinsed in PBS and immersed for 3 days in the hydrogel monomer consisting of 4% acrylamide, 0.25% VA044 Wako thermal initiator. Tubes were flushed with nitrogen gas and tissues were polymerized in a 37-°C water bath for 3 h. Active clearing was achieved using the X-Clarity Tissue clearing system (Logos Biosystems). Fluorescence imaging of CLARITY samples in Histodenz (Sigma-Aldrich, D22158) was performed using a light-sheet fluorescence microscope (Carl Zeiss LSFM Z.1) with a Fluar 4× objective lens. Images were reconstructed in 3D using TeraStitcher and Imaris software.

Clustering method. Clustering methods allowed identification of renouncing and persevering mice. Clustering was performed on the entire set of behavioural data, namely the number of active and inactive lever presses, the time of the last laser stimulation, the time until the end of the session and finally the delays between the active lever presses (restricted to the third and fourth punished sessions displaying a clear bimodal distribution in Fig. 1c). Prior to the clustering, we applied a nonlinear dimension reduction algorithm (t-distributed stochastic neighbor embedding[50]) to end up with two relevant projected variables. We then used a hierarchical clustering algorithm (with the Minkowski metric and average linkage method, using the MATLAB 'pdist', 'linkage' and 'cluster' functions) to divide the mice into two clusters based on their overall behavioural similarities (see Extended Data Fig. 1a). The mean silhouette value (computed with the MATLAB 'silhouette' function) was equal to 0.95, indicating that the clustering solution is appropriate. Mapping the perseverance onto these two clusters supported our initial separation based on visual inspection (see). To construct the ellipse around a cluster, we first fixed the centre as the mean position of the cluster. We then computed from these coordinates a covariance matrix, which was rescaled to ensure that at the end the ellipse enclosed around 98% of the points in the cluster. We got the coordinates of

the ellipse by applying the eigenvectors and eigenvalues of the covariance matrix to the unit circle (that is, for the rotation and scaling), which was finally translated by adding the mean position of the cluster initially computed.

Statistics. Samples sizes were predetermined using a power and sample size calculator. The experiments were not randomized. The investigators were blinded to genotypes during experiments and outcome assessment. Multiple comparisons were first subject to mixed-factor ANOVA defining both between- and/or within-group factors. For comparisons in which significant main effects or interaction terms were found ($P < 0.05$), further comparisons were made by a two-tailed Student's t-test with Bonferroni corrections applied when appropriate (that is, the level of significance was 0.05 adjusted by the number of comparisons). Single comparisons of between- or within-group measures were made by two-tailed non-paired or paired Student's t-test, respectively.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The dataset is available from https://doi.org/10.5281/zenodo.1474531.

45. Turiault, M. et al. Analysis of dopamine transporter gene expression pattern — generation of DAT-iCre transgenic mice. *FEBS J.* **274**, 3568–3577 (2007).
46. Lerner, T. N. et al. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
47. da Silva, J. A., Tecuapetla, F., Paixão, V. & Costa, R. M. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554**, 244–248 (2018).
48. Li, Y. et al. Serotonin neurons in the dorsal raphe nucleus encode reward signals. *Nat. Commun.* **7**, 10503 (2016).
49. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
50. Van Der Maaten, L. & Hinton, G. H. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
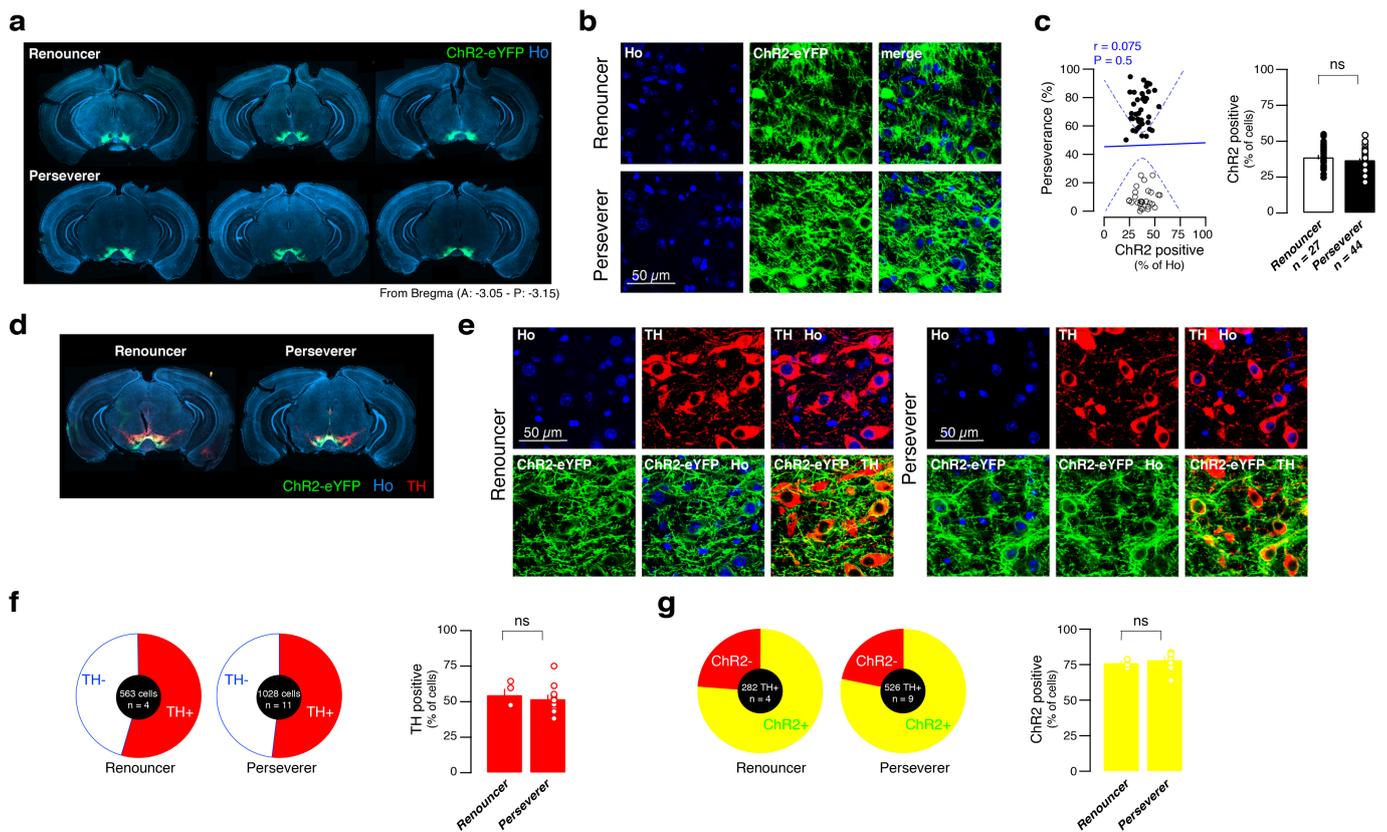
Extended Data Fig. 1 | No correlation between perseverance and VTA infection. a, Serial coronal sections of a renouncing and a persevering mouse centred on the VTA infected with AAV5-EF1a-DIO-ChR2-eYFP and nuclear Hoechst staining (Ho). We infected 109 mice and took coronal images of brains from 71 mice. **b**, High-magnification images of VTA. **c**, oDASS perseverance as a function of the infection rate and group data ($n = 71$ mice, Pearson's $r = 0.075$). Infection rate was determined as the number of ChR2–eYFP-positive cells normalized to the total number of cells based on Hoechst staining. Note that mice from which sagittal sections were obtained are not included in this quantification. **d**, Coronal sections of a renouncing and a persevering mouse that show VTA as above. Sections were additionally stained for tyrosine hydroxylase (TH) using a Cy3-conjugated secondary antibody. Staining was performed in slices from 15 mice. **e**, High-magnification images of VTA. **f**, Quantification of TH-positive neurons in the VTA from a subset of renouncing and persevering mice. **g**, Percentage of TH neurons infected with ChR2. Data are mean $\pm$ s.e.m. See Supplementary Table 1 for complete statistics.
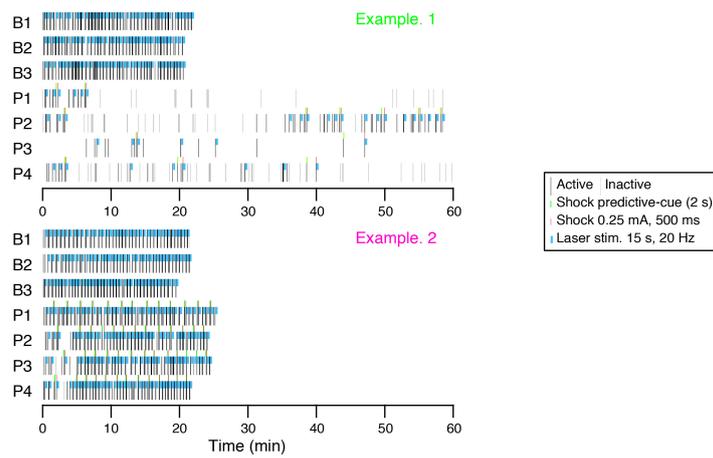
Extended Data Fig. 2 | Examples from two mice of lever presses and outcomes during oDASS. Raster plots for two mice, one keeping a high stimulation rate during punished sessions (example 2) and one falling to a low stimulation rate during punished sessions (example 1). Every action and the associated outcome as a function of time is shown for three baseline (B) and four punished (P) sessions.
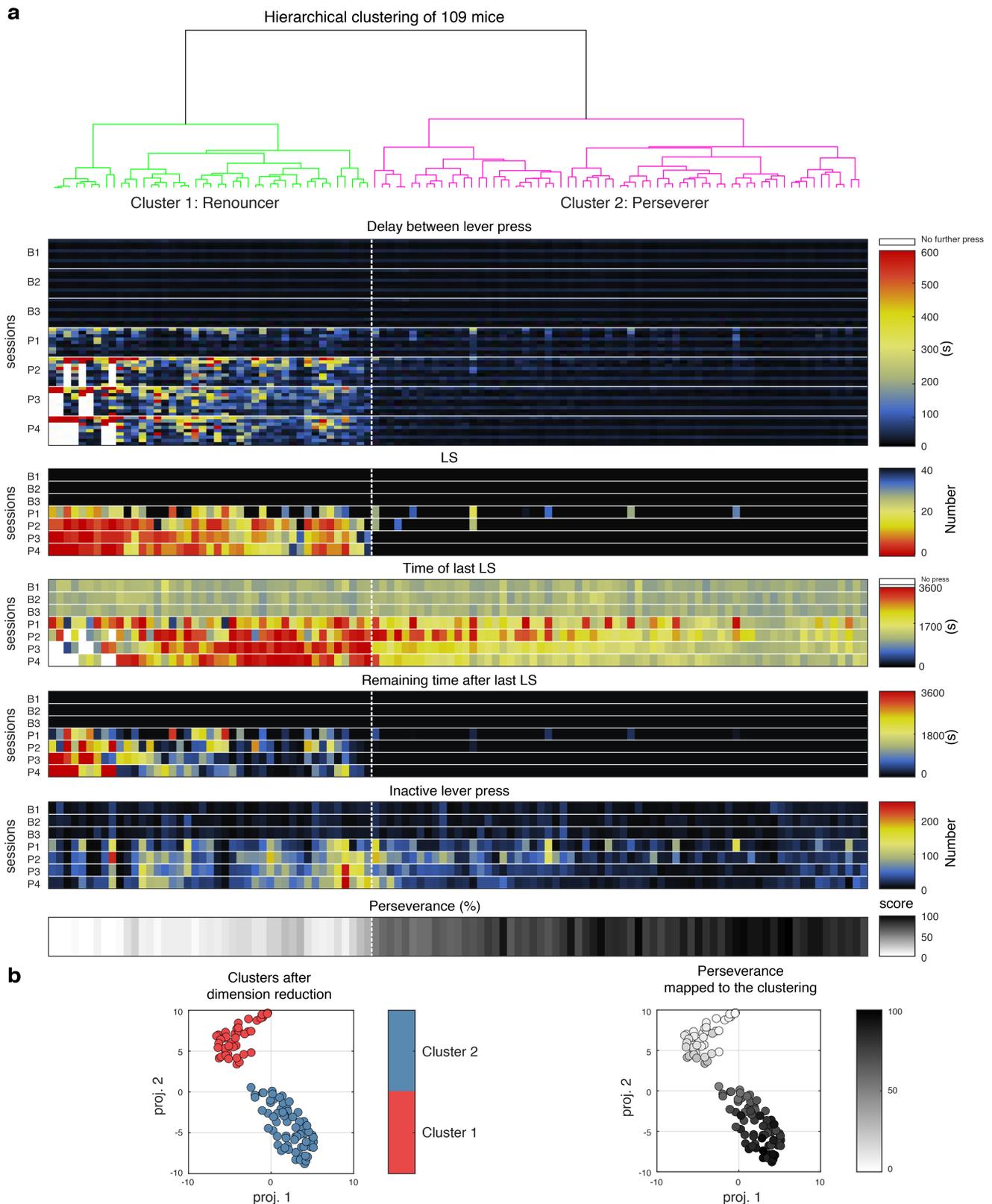
Extended Data Fig. 3 | Emergence of two clusters of mice with punished oDASS. a, Hierarchical clustering of the entire dataset (time-event for eight parameters during three baseline and four punished sessions). Each column corresponds to a mouse. Heat maps for delays between the active lever presses, the number of laser stimulation events, the time of the last laser stimulation, the time remaining in a trial and the number of inactive lever press in one mouse are plotted. Two clusters are found in the resulting dendrogram (green, renouncer; pink, perseverer). Vertical dashed lines separate renouncing from persevering mice. Heat map (greyscale) represents the perseverance (measured as the oDASS rate during punished sessions 3 and 4, normalized to the baseline sessions) of each mouse as a function of the clustering. **b**, Before clustering, we applied a nonlinear dimension reduction to project the high-dimensional dataset into a two-dimensional representation (left). Mapping the perseverance onto this map (right) shows that this variable can be used to categorize the mice as renouncing and persevering mice.
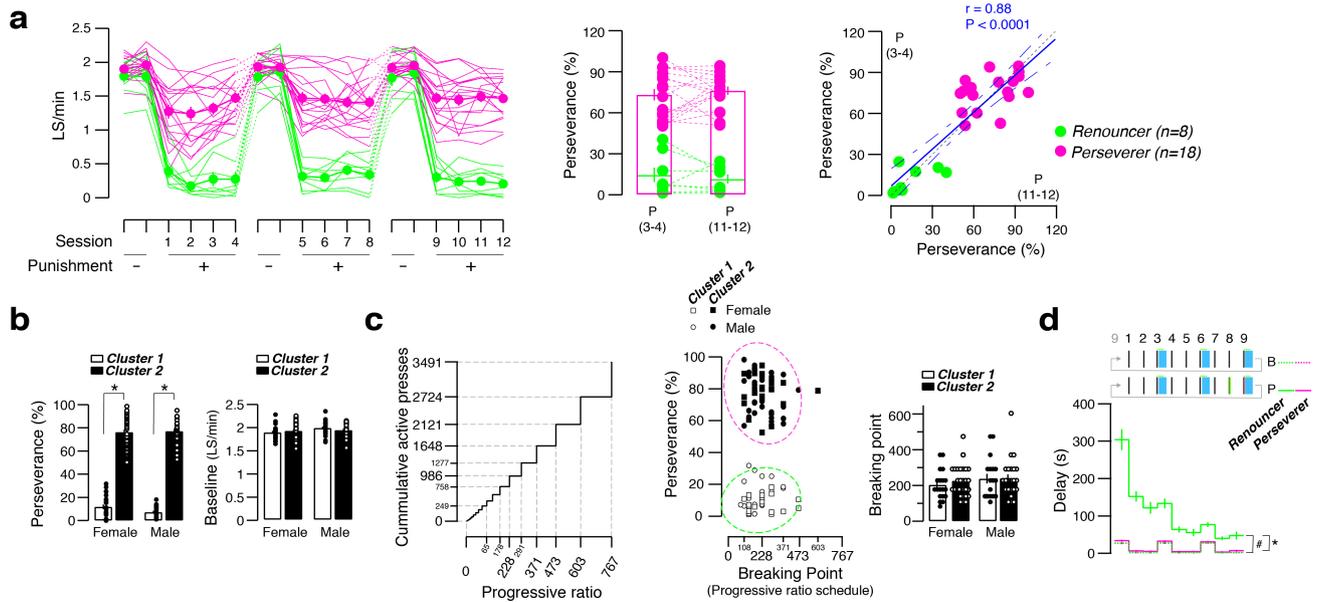
Extended Data Fig. 4 | Stability of perseverance and oDASS acquisition parameters. a, Three blocks of two baseline and four punished oDASS sessions were performed during a two-month period. Left, oDASS rate as a function of advancing sessions and average for the two groups. Middle, perseverance measured during punishment sessions 3–4 was compared to sessions 11–12. Perseverance was calculated as the average oDASS during the two punished sessions normalized to the corresponding baseline rate. No effect of punishment block was detected. Right, correlation between perseverance between first and last block (sessions 3–4 versus session 11–12), Pearson's $r = 0.88$ ($n = 26$). **b,** Left, perseverance and baseline rate for male and female mice of the two clusters (for perseverance, ANOVA followed two-sided t-test: *$P < 0.0001$, $t_{60} = 23.02$ for persevering versus renouncing female mice ($n = 39$ and 23 mice, respectively); *$P < 0.0001$, $t_{45} = 22.30$ for persevering versus renouncing male mice ($n = 20$ and 27 mice, respectively). Right, no difference in baseline rate was detected between the two clusters, nor between male and female mice

($n = 109$ mice). **c,** Left, cumulative active presses as a function of the progressive ratio. Middle, perseverance as a function of the breakpoint during the progressive ratio schedule for male (42, squares) and female (56, circles). Note that in Fig. 1e, data are presented with cumulative active presses, which avoids the steps that are observed inherent to the breakpoint plot. No difference in breakpoint was detected between the two clusters, nor between male and female mice ($n = 98$ mice). **d,** Temporal structure of an oDASS trial showing the delays between active press during baseline and punished sessions for cluster 1 and cluster 2 mice (renouncing and persevering mice, respectively). Delays were increased in renouncing mice (ANOVA followed by two-sided t-test: *$P < 0.05$ in punished sessions for persevering versus renouncing mice for every delay ($n = 66$ and 43 mice, respectively); #$P < 0.05$; baseline/ punished for renouncing mice for every delay). Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.
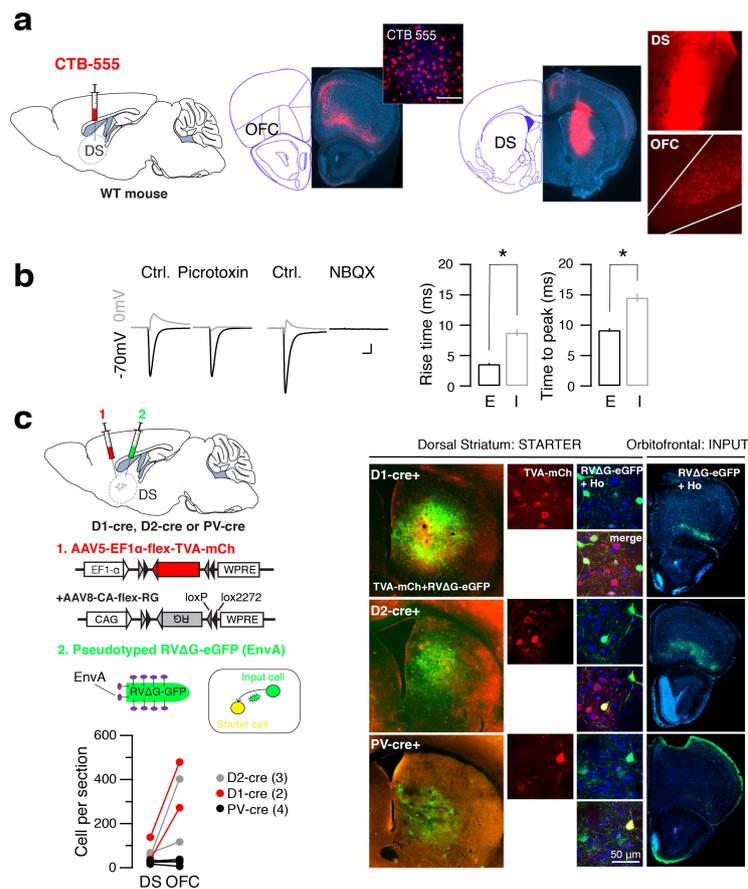
Extended Data Fig. 5 | Characterization of the connection between OFC and striatum. a, Retrograde tracing from the striatum with cholera toxin subunit B coupled to a red dye (CTB-555). Scale bar, 50 μm. Experiment was repeated in $n = 8$ mice. **b**, OFC–striatum optogenetic stimulation and recordings of EPSCs, blocked by the AMPAR antagonist (NBQX) and IPSCs blocked by GABA$_A$ antagonist (picrotoxin). Rise time and time to peak for IPSCs and EPSCs indicate a feed-forward circuit between principal neurons of the OFC and striatum (two-sided paired t-test for IPSCs versus EPSCs: $t_{25} = 9.55$, $*P < 0.0001$ and $t_{25} = 7.57$, $P < 0.0001$ for rise time and time to peak, respectively, $n = 26$ cells from five mice). Scale bars, 20 ms, 100 pA. **c**, Schematic of retrograde tracing from specific cell types of the striatum using a rabies virus that was injected in transgenic Cre-mouse lines encoding the dopamine D1 or D2 receptor and parvalbumin. A first injection (red) leads to cell-type-specific expression of the EnvA receptor TVA and the RG protein. After two weeks an EnvA-pseudotyped and glycoprotein (ΔG)-deleted rabies virus (EnvA and RVΔG–GFP) is injected (green) and taken up by the cells that express TVA and thus turn yellow (starters). Trans-complemented with glycoprotein by infection of the AAV8-CA-flex-RG and RVΔG–GFP transsynaptically caused a spread to upstream neurons (inputs). The injection site in the striatum (left) and high-magnification images show starter cells. Retrogradely infected neurons in the OFC at low and higher magnification (right). See Supplementary Table 1 for complete statistics. **a**, **c**, Images reproduced from Paxinos and Franklin[44], copyright © 2001.
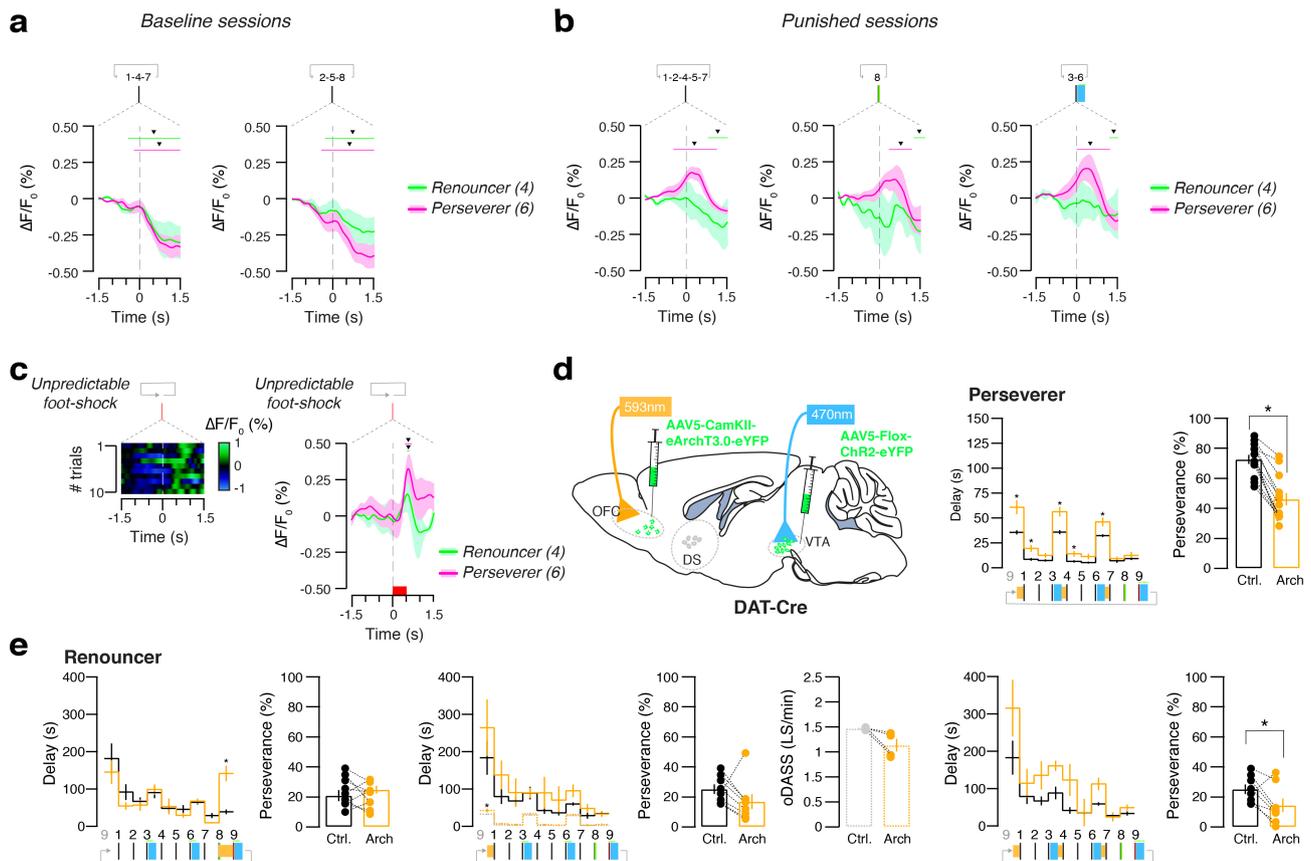
Extended Data Fig. 6 | Activity and manipulation of OFC–striatum projection during oDASS. a, Calcium signal ($\Delta F/F_0$, mean $\pm$ s.e.m.) around active press 1 and 2 of the FR3 schedule (numbers 1, 4, 7 and 2, 5, 8) during a baseline session for renouncing and persevering mice (green diamonds and pink bars indicate a significant deviation from baseline, $n = 4$ and 6 mice). **b**, Averaged calcium signal ($\Delta F/F_0$, mean $\pm$ s.e.m.) around active press 1 and 2 of the FR3 schedule (all but press number 8), around the active press number 8 (leading to the shock-associated cue) and around the lever press that terminates the non-shock FRs in punished sessions for renouncing and persevering mice (green diamonds and pink bars indicate a significant deviation from baseline, $n = 4$ and 6 mice). **c**, Example of trial activity map of the calcium signal ($\Delta F/F_0$, mean $\pm$ s.e.m.) around an unpredictable foot shock (500 ms, 0.25 mA, repeated in 10 times in one mouse). For each animal, 10 unpredictable foot shocks were delivered during a separate recording. Grouped data for the calcium signal ($\Delta F/F_0$, mean $\pm$ s.e.m.) around an unpredictable foot shock for renouncing and persevering mice (green diamonds and pink bars indicate a significant deviation from baseline, $n = 4$ and 6 mice). See Supplementary Table 1 for statistics. **d**, Scheme of a mouse brain infected with eArchT3.0–eYFP in the OFC and with ChR2–eYFP in the VTA (left). For persevering mice, OFC inhibition with eArchT3.0

between oDASS and the next FR initiation (or for a maximum of 90 s) delayed the next press of persevering mice (ANOVA followed by two-sided t-test: *$P < 0.05$ when comparing control versus eArchT3.0 delays during punished sessions, $n = 13$ mice). Perseverance changed (from 73% to 46%) as a consequence of eArchT3.0 stimulation (two-sided paired t-test: $t_{12} = 9.13$, *$P < 0.0001$, $n = 13$ mice for control versus eArchT3.0 before each FR initiation). **e**, For renouncing mice, OFC inhibition with eArchT3.0, after the punishment-predictive cue, between punished oDASS and the next FR initiation (or for a maximum of 90 s) or between each oDASS and the next FR initiation slightly delayed the next press (ANOVA followed two-sided paired t-test: *$P < 0.05$ when comparing control and eArchT3.0 delays during punished sessions). Perseverance was reduced as a consequence of eArchT3.0 stimulation between each oDASS and the next FR initiation (two-sided paired t-test: $t_7 = 2.62$, *$P = 0.034$, $n = 8$ mice for control versus eArchT3.0 before each FR initiation). The oDASS rate during a baseline session was not significantly changed between punished oDASS and the next FR initiation (or for a maximum of 90 s) by inhibition with eArchT3.0. Data are mean $\pm$ s.e.m. See Supplementary Table 1 for complete statistics. **d**, Line drawing modified from Paxinos and Franklin[44], copyright © 2007.
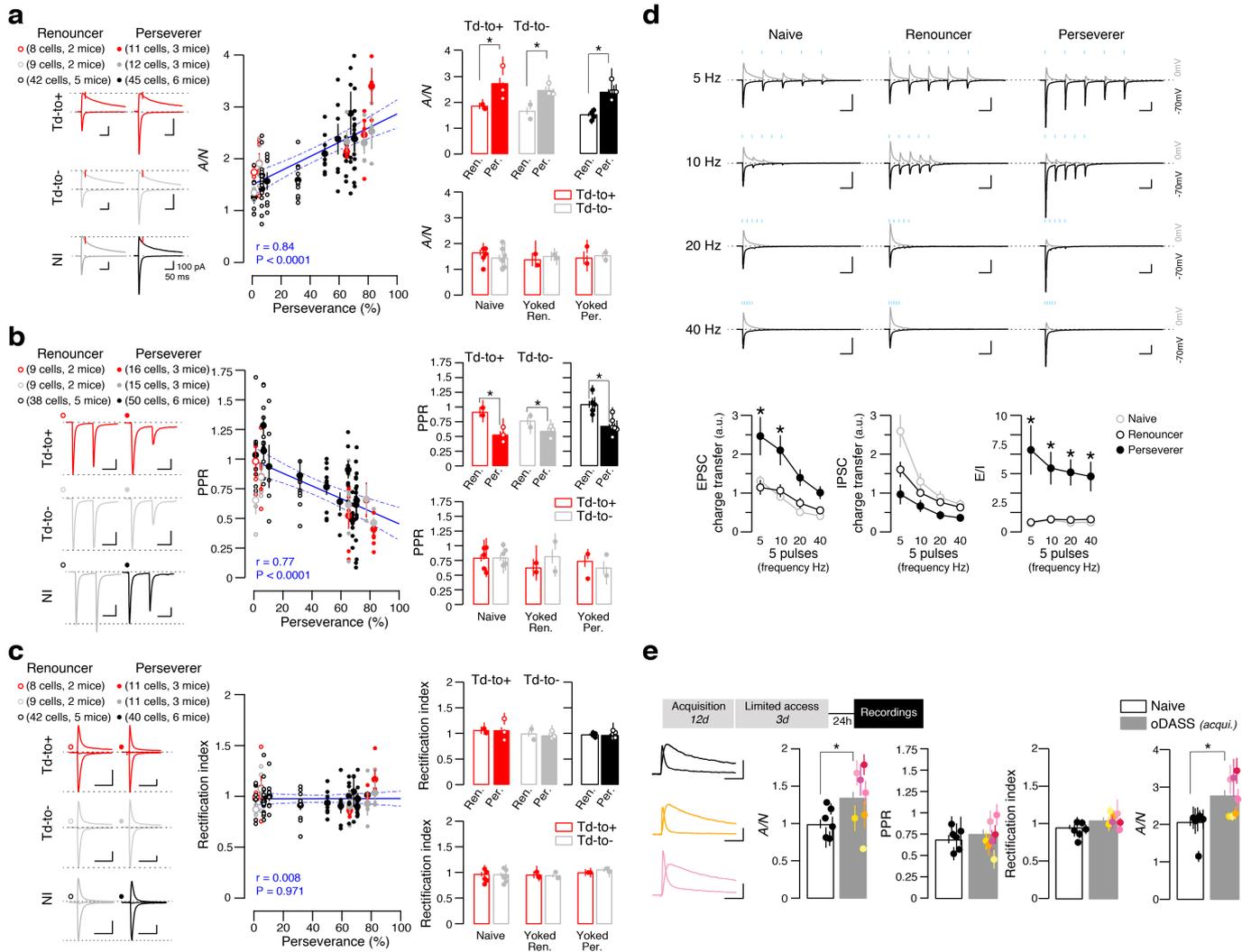
Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | Plasticity at OFC–striatum synapses correlates with compulsive oDASS. a**, Average of 10 sweeps for AMPAR EPSCs recorded at −70 mV and for EPSCs recorded at +40 mV in slices from renouncing and persevering mice, with or without SPN identification with tdTomato. NMDA amplitude was analysed 20 ms after the peak of the EPSC recorded at +40 mV. The AMPAR/NMDAR ratio for every recorded neuron as a function of perseverance, data are mean ± s.e.m. per animal and correlation (Pearson's $r = 0.84$, $P < 0.0001$; $n = 127$ cells from 16 mice). Data are mean AMPAR/NMDAR ratio for renouncing and persevering mice (ANOVA followed by two-sided t-test: $t_{17} = 3.13$, $*P = 0.007$ and $t_{19} = 3.53$; $*P = 0.002$ for tdTomato$^+$ and tdTomato$^-$, respectively ($n = 8$ and 11 mice; 9 and 12 cells); $t_{85} = 6.68$, $*P < 0.0001$, renouncing versus persevering for non-identified neurons (42 and 45 cells, respectively)). Data are mean AMPAR/NMDAR ratio for naive mice and for mice yoked to renouncing or persevering mice, with Drd1a–tdTomato identification (naive mice, $n = 14$ tdTomato$^+$ cells from 5 mice and $n = 12$ tdTomato$^-$ cells from 7 mice; mice yoked to renouncing mice, $n = 6$ tdTomato$^+$ cells from 2 mice and $n = 9$ tdTomato$^-$ cells from 2 mice; mice yoked to persevering mice, $n = 6$ tdTomato$^+$ cells from 2 mice and $n = 8$ tdTomato$^-$ cells from 2 mice). **b**, Average of 10 sweeps for AMPAR EPSCs recorded at −70 mV, with two short light pulses spaced with a 76-ms interval in slices from renouncing and persevering mice, with or without SPN identification using tdTomato. PPR is the ratio of the amplitudes of the second EPSC over the first. PPR for every recorded neuron as a function of perseverance, mean ± s.e.m. per animal and correlation (Pearson's $r = -0.77$, $P < 0.0001$; $n = 137$ cells from 16 mice). Data are mean PPR for renouncing and persevering mice (ANOVA followed by two-sided t-test: $t_{23} = 3.97$, $*P = 0.0005$ and $t_{22} = 1.72$, $P = 0.183$ for tdTomato$^+$ and tdTomato$^-$, respectively ($n = 9$ and 15 mice; 9 and 16 cells); $t_{86} = 5.66$, $*P < 0.0001$, renouncing versus persevering for non-identified neurons ($n = 38$ and 50 cells, respectively)). Data are mean PPR for naive mice and for mice yoked to renouncing or persevering mice, with Drd1a–tdTomato identification (naive mice, $n = 13$ tdTomato$^+$ cells from 5 mice and $n = 11$ tdTomato$^-$ cells from 7 mice; mice yoked to renouncing mice, $n = 6$ tdTomato$^+$ cells from 2 mice and $n = 10$ tdTomato$^-$ cells from 2 mice; mice yoked to persevering mice, $n = 7$ tdTomato$^+$ cells from 2 mice and $n = 8$ tdTomato$^-$ cells from 2 mice). **c**, Left, average of 10 sweeps for AMPAR EPSCs recorded at −70, 0 and +40 mV in slices from renouncing and persevering mice, with or without SPN identification using tdTomato. The rectification of the AMPAR EPSCs was calculated as the ratio of the chord conductance calculated at negative potential divided by chord conductance at positive

potential. Middle, rectification index for every recorded neuron as a function of perseverance, mean ± s.e.m. per animal and correlation (Pearson's $r = 0.008$, $P = 0.971$; $n = 121$ cells from 16 mice). Right, rectification index for renouncing and persevering in tdTomato$^+$ or tdTomato$^-$ cells (8 and 11 mice; 9 and 11 cells) and in unidentified SPNs (42 and 40 cells, respectively). Data are mean rectification index for naive mice and for mice yoked to renouncing or persevering mice with Drd1a–tdTomato identification (naive mice, $n = 13$ tdTomato$^+$ cells from 5 mice and $n = 11$ tdTomato$^-$ cells from 7 mice; mice yoked to renouncing mice, $n = 6$ tdTomato$^+$ cells from 2 mice and $n = 9$ tdTomato$^-$ cells from 2 mice; mice yoked to persevering mice, $n = 7$ tdTomato$^+$ cells from 2 mice and $n = 8$ tdTomato$^-$ cells from 2 mice). **d**, Average of 10 sweeps for EPSCs recorded at −70 mV and IPSCs recorded at 0 mV in slices from renouncing and persevering mice and in slices from naive mice. Five pulses were given at different frequencies (5, 10, 20 and 40 Hz) and the charge transfer was measured (area under the curve). The excitatory/inhibitory ratio (E/I) was calculated as the ratio charge transfer for EPSCs over IPSCs. The charge transfer of EPSCs was higher in slices from persevering mice at low frequencies (ANOVA followed by two-sided t-test: $t_{27} = 4.75$, $*P < 0.0001$; $t_{27} = 3.75$, $*P = 0.0007$; $t_{27} = 2.37$, $P = 0.057$; $t_{27} = 1.64$, $P = 0.306$ for 5, 10, 20 and 40 Hz, respectively (17 and 12 cells)). The charge transfer of IPSCs was not different between persevering and renouncing mice (17 and 12 cells, respectively). The ratio of charge transfer for EPSCs over IPSCs was higher in slices from persevering mice (ANOVA followed by two-sided t-test: $t_{27} = 6.22$, $*P < 0.0001$; $t_{27} = 4.39$, $*P < 0.0001$; $t_{27} = 4.07$, $P = 0.0002$; $t_{27} = 3.67$, $P = 0.001$ for 5, 10, 20 and 40 Hz, respectively (17 and 12 cells)). Measurements were obtained from four renouncing mice, three persevering mice and three naive mice. **e**, Average of 10 sweeps for AMPAR EPSCs in the presence of D-AP5 (50 μM) and NMDAR EPSCs isolated by subtraction for oDASS mice and for naive mice. Mean AMPAR/NMDAR ratio, PPR and rectification index for naive and oDASS mice ($*P < 0.05$ for t-test comparing naive/oDASS). Each dot represents the mean ± s.e.m. for all cells obtained in a given mouse. Recordings were obtained from six naive mice and seven oDASS mice (PPR: 48 cells from oDASS mice compared to 29 cells from naive mice; rectification index: 52 cells from oDASS mice compared to 24 cells from naive mice; AMPAR/NMDAR ratio with pharmacological isolation: 42 cells from oDASS mice compared to 23 cells from naive mice; AMPAR/NMDAR ratio without pharmacological isolation: 52 cells from oDASS mice compared to 26 cells from naive mice). Scale bars, 200 pA, 50 ms. Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.
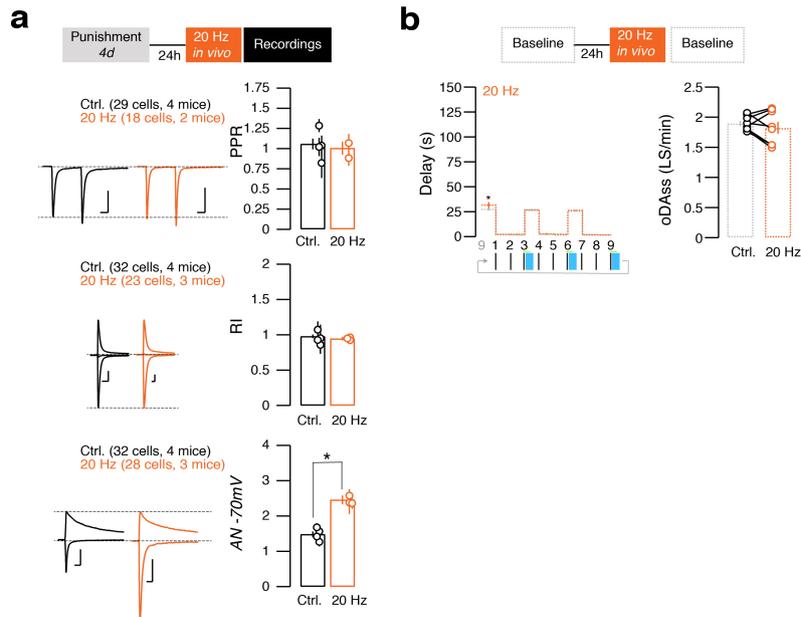
Extended Data Fig. 8 | Synaptic properties in persevering mice after 20-Hz stimulation in vivo. a, Average traces for EPSCs recorded to determine PPR, rectification index (RI) and AMPAR/NMDAR ratio without pharmacological isolation. Ex vivo measurement of the PPR, rectification index and AMPAR/NMDAR ratio after in vivo stimulation of OFC–striatum terminals at 20 Hz for 1 min in renouncing mice (for PPR: control versus 20 Hz, two-sided t-test: $t_{45} = 0.48$, $P = 0.63$, $n = 29$ and 18 cells, respectively; for rectification index: control versus 20 Hz: $t_{53} = 0.90$, $P = 0.37$, $n = 32$ and 23 cells respectively; for AMPAR/

NMDAR ratio: control versus 20 Hz: $t_{58} = 6.79$, *$P < 0.0001$, $n = 32$ and 28 cells, respectively). **b**, Effect of 20-Hz stimulation of OFC–striatum prior to a baseline session ($n = 8$ mice). Delay to engage the next action was not changed (ANOVA followed by two-sided t-test: *$P < 0.05$ when comparing control and 20 Hz). oDASS rate was not modified by 20 Hz stimulation before a baseline session ($t_7 = 0.97$, $P = 0.36$, $n = 8$ mice for control versus 20 Hz before a baseline session). Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.
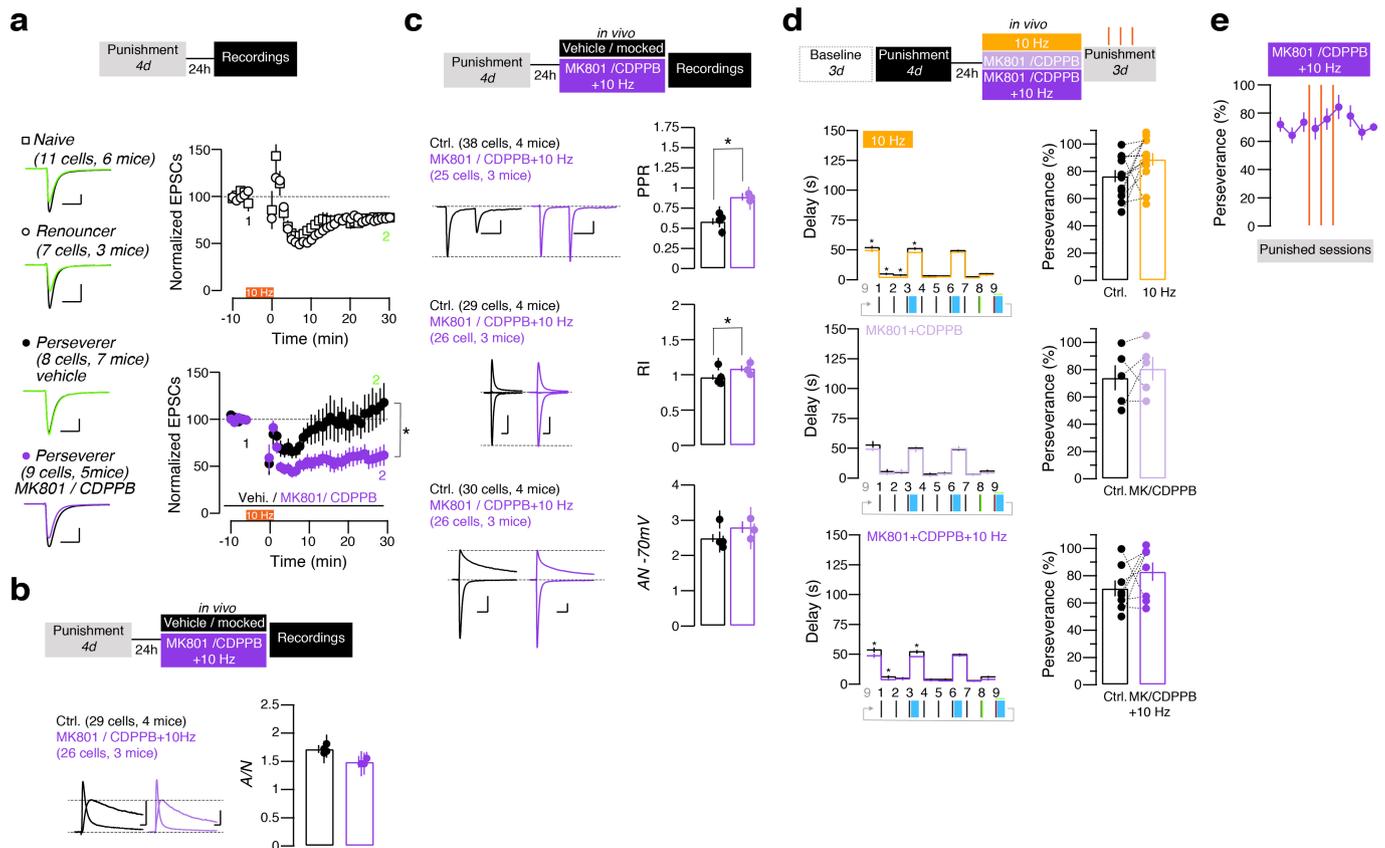
Extended Data Fig. 9 | Absence of an effect on compulsive behaviour by normalization of release probability in persevering mice.
a, Average traces for EPSCs recorded immediately before and 30 min after the LTD protocol (10 Hz for 5 min) and grouped data. In slices from persevering mice, LTD is unmasked by bath application of mGluR5 PAM (CDPPB, 100 μM) and MK801 (NMDAR blocker, 10 μM) (two-sided t-test: $t_{15} = 2.89$, *$P = 0.037$, $n = 8$ and 9 cells, respectively).
b, AMPAR/NMDAR ratio was left unchanged by in vivo stimulation of OFC–striatum terminals with 10 Hz in presence of MK801 and CDPPB (0.3 and 30 mg kg^{-1}, respectively) in persevering mice (two-sided t-test: $t_{53} = 1.86$, $P = 0.069$ for control versus 10 Hz treated with MK801 and CDPPB ($n = 29$ and 26 cells, respectively)). **c**, PPR was normalized by in vivo stimulation of OFC–striatum terminals with 10 Hz in the presence of MK801 and CDPPB in persevering mice (two-sided t-test: $t_{61} = 4.94$, $P < 0.0001$ for control versus 10 Hz with MK801 and CDPPB ($n = 38$ and 25 cells, respectively)). The rectification index was different between controls and mice treated in vivo with the stimulation of OFC–striatum terminals at 10 Hz in the presence of MK801 and CDPPB (two-sided t-test:

$t_{53} = 2.25$, *$P = 0.029$ for control versus 10 Hz with MK801 and CDPPB ($n = 29$ and 26 cells, respectively)). AMPAR/NMDAR ratio without pharmacological isolation was not different between controls and mice treated in vivo with the stimulation of OFC–striatum terminals at 10 Hz in the presence of MK801 and CDPPB (two-sided t-test: $t_{54} = 1.57$, $P = 0.12$ for control versus 10 Hz with MK801 and CDPPB ($n = 30$ and 26 cells, respectively)). **d**, Plots for delay between lever presses in punished sessions 12 h after 10 Hz, MK801 and CDPPB or 10 Hz with MK801and CDPPB (ANOVA followed by two-sided paired t-test: *$P < 0.05$ when comparing control/treatment delays during punished sessions). Perseverance is not modified by any treatment (two-sided t-test: $t_{12} = 2.12$, $P = 0.056$, $n = 13$ mice for control versus 10 Hz; $t_4 = 0.73$, $P = 0.51$, $n = 5$ mice for control versus MK801 and CDPPB; $t_7 = 1.31$, $P = 0.231$, $n = 8$ mice for control versus 10 Hz with MK801 and CDPPB). **e**, During additional punished sessions without renewal of the intervention, perseverance remained unchanged ($n = 8$ mice). Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.
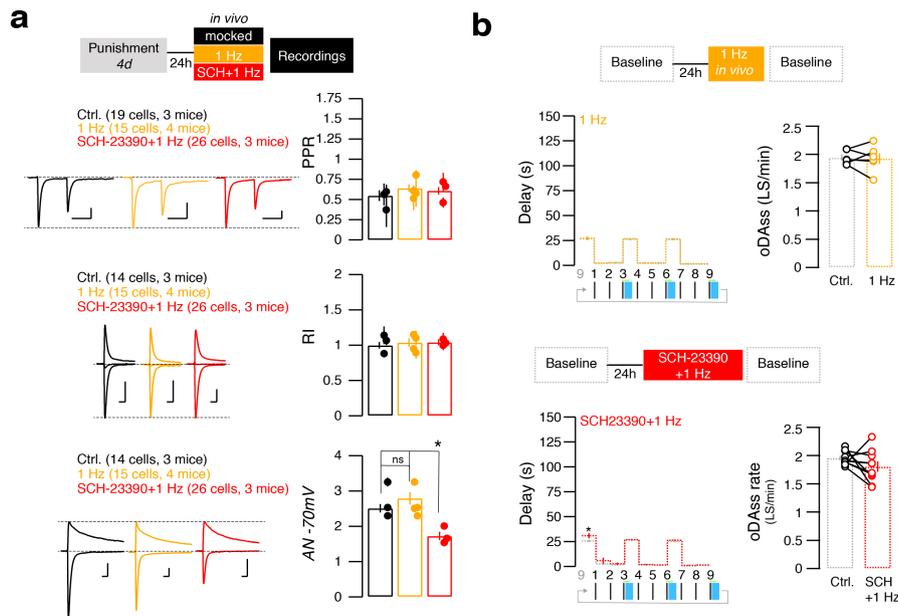
Extended Data Fig. 10 | Effects of SCH23390 and 1 Hz in vivo in persevering mice. a, Average traces of EPSCs recorded to determine PPR, rectification index and AMPAR/NMDAR ratio without pharmacological isolation. Ex vivo measurement of the PPR, rectification index and AMPAR/NMDAR ratio after in vivo stimulation of OFC–striatum terminals at 1 Hz for 5 min in persevering mice, in the presence or absence of SCH23390 (ANOVA followed by two-sided t-test: for PPR: $t_{32} = 1.10$, $P = 0.83$; $t_{39} = 0.64$, $P > 0.99$; and $t_{43} = 0.57$, $P > 0.99$ for control versus 1 Hz ($n = 19$ and 15 cells, respectively); 1 Hz versus 1 Hz with SCH23390 ($n = 15$ and 26 cells, respectively); and control versus 1 Hz with SCH23390 ($n = 19$ and 26 cells, respectively); for rectification index: $t_{27} = 0.74$, $P > 0.99$; $t_{39} = 0.17$, $P > 0.99$; and $t_{38} = 1.00$, $P = 0.97$ for control versus 1 Hz ($n = 14$ and 15 cells, respectively); 1 Hz versus 1 Hz with SCH23390 ($n = 15$ and 26 cells, respectively); and control versus 1 Hz with SCH23390

($n = 14$ and 26 cells, respectively); for AMPAR/NMDAR ratio: $t_{27} = 1.64$, $P = 0.32$; $t_{39} = 5.93$, $P < 0.0001$; and $t_{38} = 3.96$, $P = 0.0007$ for control versus 1 Hz ($n = 14$ and 15 cells, respectively); 1 Hz versus 1 Hz with SCH23390 ($n = 15$ and 26 cells, respectively); and control versus 1 Hz with SCH23390 ($n = 14$ and 26 cells, respectively)). **b**, Delay to engage the next action was not changed (ANOVA followed by two-sided paired t-test: *$P < 0.05$ when comparing control versus 1 Hz or control versus 1 Hz with SCH23390, $n = 10$ mice). oDASS rate was not modified by 1 Hz or 1 Hz with SCH23390 prior to a baseline session ($t_5 = 0.22$, $P = 0.84$, $n = 6$ mice for control versus 1 Hz before a baseline session and $t_9 = 1.48$, $P = 0.17$, $n = 10$ mice for control versus 1 Hz with SCH233₉0 before a baseline session). Data are mean ± s.e.m. See Supplementary Table 1 for complete statistics.

nature research

Corresponding author(s): Christian Lüscher

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Igor Pro 7 (Wave metrics Electrophysiology); MED-PC IV 4.34, Med Associate (Operant behaviour); TDT synapse version 84 Tucker-Davis (photometry)
Data analysis	Igor Pro 7; Ms-Excel 14.4.4; GraphPad Prism 7.02

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Upon publication the raw data will be available at "https://archive-ouverte.unige.ch"

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size calculators such as powerandsamplesize.com. For the clustering algorithm a Monte Carlo simulation was performed.
Data exclusions	We exluded the follwoing data: Voltage clamp recordings where access resistance varied more than 15% Current clamp recordings at i=0 when resting membrane potential varied more than 10%. Animals where post-hoc verification of the injection site failed.
Replication	A detailed experimental protocol is provided to facilitate replication by others. For all electrophysiology experiments we used interleaved controls, between experimenters (n=3) reproduction. For behavioral experiments the details are provided in the legends and statistical
Randomization	Permutation of treatment order (e.g. application of pharmacological agents and optogenetic treatment protocols)
Blinding	For ex vivo patch recordings, experimenter was blind as to the behavioral phenotype

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Rabbit polyclonal anti-Tyrosine Hydroxylase (MILLIPORE AB152, polyclonal, lot 2722866, 500x); Donkey anti-Rabbit CY3, (Millipore AP182C, polyclonal, lot number 2397069, 500x diluted)
Validation	Anti-Tyrosine Hydroxylase Antibody for TH and has been previously validated for use in ELISA, IF, IH, IH(P), IP and WB in mice according to the manufacturers website.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The following mouse lines were used: DAT-cre (from Günter Schultz, DKFZ Heidelberg); DAT B6.SJL-Slc6a3tm1.1(cre)Bkmn/J known as DAT-IRES-Cre crossed Drd1a-tdTomato from Jackson Laboratories; Tg(Drd1-cre)120Mxu; Tg(Drd2-Cre)ER43Gsat and 129P2-Pvalbtm1(cre)Arbr/J from CHARLES RIVER. For all line we used both male and female mice, aged 8-24 weeks.
Wild animals	Not applicable
Field-collected samples	Not applicable