Article

Cell-type-specific asynchronous modulation of PKA by dopamine in learning

https://doi.org/10.1038/s41586-020-03050-5

Received: 11 November 2019

Accepted: 30 October 2020

Published online: 23 December 2020

Check for updates

Suk Joon Lee[1], Bart Lodder[1], Yao Chen[1,2], Tommaso Patriarchi[3,4], Lin Tian[3] & Bernardo L. Sabatini[1]✉

Reinforcement learning models postulate that neurons that release dopamine encode information about action and action outcome, and provide a teaching signal to striatal spiny projection neurons in the form of dopamine release[1]. Dopamine is thought to guide learning via dynamic and differential modulation of protein kinase A (PKA) in each class of spiny projection neuron[2]. However, the real-time relationship between dopamine and PKA in spiny projection neurons remains untested in behaving animals. Here we monitor the activity of dopamine-releasing neurons, extracellular levels of dopamine and net PKA activity in spiny projection neurons in the nucleus accumbens of mice during learning. We find positive and negative modulation of dopamine that evolves across training and is both necessary and sufficient to explain concurrent fluctuations in the PKA activity of spiny projection neurons. Modulations of PKA in spiny projection neurons that express type-1 and type-2 dopamine receptors are dichotomous, such that these neurons are selectively sensitive to increases and decreases, respectively, in dopamine that occur at different phases of learning. Thus, PKA-dependent pathways in each class of spiny projection neuron are asynchronously engaged by positive or negative dopamine signals during learning.

Across phylogeny, dopamine (DA) release in the brain induces cellular plasticity that promotes behavioural adaptation[3]. In mammals, DA action in the nucleus accumbens (NAc)—a striatal region that is heavily innervated by DA-releasing neurons (DANs) of the ventral tegmental area (VTA)—mediates the association of motor actions with action outcomes, as necessary for individuals to learn to repeat behaviours that lead to good outcomes[1,2,4,5]. Previous manipulations of VTA DANs and DA levels in the NAc have established the sufficiency of DA release in the NAc for action reinforcement[6–9]. Furthermore, the activity of VTA DANs and DA levels in the NAc encode reward prediction error (the difference between the actual and expected values of the outcome of an action)[10–13].

The anatomical and molecular organization of spiny projection neurons (SPNs), the principal cells of the NAc, suggest an antagonistic model of DA action on SPNs[2,14]. SPNs of the NAc (analogous to the direct- and indirect-pathway SPNs of the dorsal striatum) consist of striatomesencephalic SPNs (which innervate midbrain regions) and striatopallidal SPNs (which innervate the ventral pallidum)[15,16]. This anatomic division correlates with molecular differences: striatomesencephalic SPNs express $G_{\alpha s}$-coupled type-1 DA receptors (D1Rs) through which DA enhances cAMP production and PKA activity, whereas striatopallidal SPNs express $G_{\alpha i/o}$-coupled type-2 DA receptors (D2Rs), which inhibit cAMP production and suppress PKA[15,16]. Models of reinforcement learning[2] incorporate these differences and link DA transients that encode reward prediction error to the PKA-dependent modulation of excitability[17,18], synaptic plasticity[19,20] and transcription[21,22] in SPNs.

Here we investigate the relationship between DA and net PKA activity in SPNs (that is, the balance between PKA and phosphatase activities) in freely behaving mice undergoing reward-based learning, using multichannel fibre photometry and fluorescence lifetime photometry (FLiP)[23]. Our results support a model of learning, dependent on DA and basal ganglia, in which PKA in SPNs that express D1R and D2R (hereafter, D1R-SPNs and D2R-SPNs, respectively) is asynchronously engaged to mediate the action of positive and negative DA signals during learning.

Dopamine receptors can regulate SPN PKA in vivo

To examine the ability of each class of DA receptor to modulate PKA in awake mice, we monitored net PKA activity in SPNs after the intraperitoneal delivery of DA receptor agonists and antagonists. We expressed fluorescence lifetime microscopy A-kinase activity reporter (FLIM-AKAR), which has previously been validated in vitro[24,25] and in vivo[23], in either D1R-SPNs or D2R-SPNs in the NAc by injecting an adeno-associated virus that expressed the sensor in a Cre-dependent manner into transgenic mice (Fig. 1a). Fluorescence was detected using an optical fibre, and fluorescence lifetimes were measured with time-correlated single-photon counting at 1 Hz (Fig. 1a, Extended Data Fig. 1).

In *Drd1a-cre* (*Drd1a* is also known as *Drd1*) mice, FLIM-AKAR expressed in D1R-SPNs of the NAc reported an increase in net PKA

[1]Howard Hughes Medical Institute, Department of Neurobiology, Harvard Medical School, Boston, MA, USA. [2]Department of Neuroscience, Washington University School of Medicine, St Louis, MO, USA. [3]Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA. [4]Institute of Pharmacology and Toxicology, University of Zurich, Zurich, Switzerland. ✉e-mail: bsabatini@hms.harvard.edu

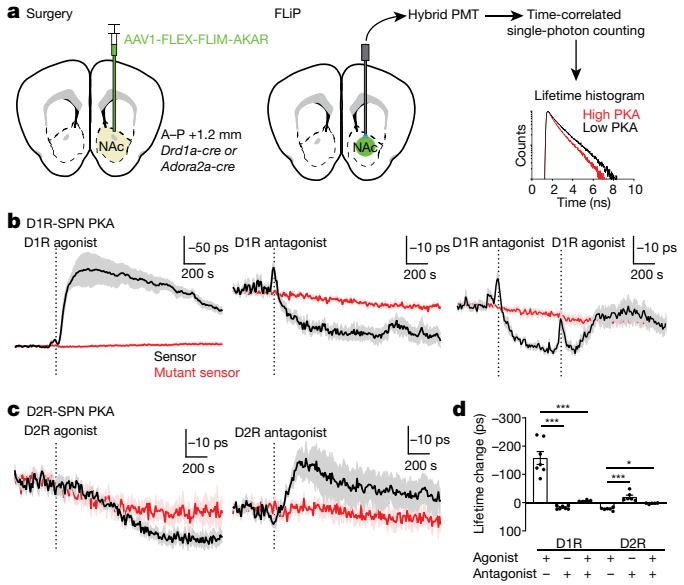


Fig. 1 | FLiP reveals bidirectional changes in PKA activity in SPNs in vivo. **a**, Left, schematic describing viral injection and optical fibre implantation. Right, hybrid photomultiplier tube (PMT) and time-correlated single-photon counting were used to measure the fluorescence lifetime of FLIM-AKAR, which shortens upon PKA phosphorylation. A–P, anterior–posterior. **b**, Average lifetime changes of FLIM-AKAR (black) and FLIM-AKAR(T391A) (red) expressed in D1R-SPNs in response to D1R agonist (left), antagonist (middle) and antagonist + agonist (right) ($n = 7$ mice). Dashed lines indicate the end of the injection. Plotted as mean ± s.e.m. across mice. **c**, Average lifetime changes of FLIM-AKAR in D2R-SPNs in response to D2R agonist (left) and antagonist (right) ($n = 6$ mice for FLIM-AKAR, 5 for FLIM-AKAR(T391A)) plotted as in **b**. **d**, Lifetime changes in FLIM-AKAR (from **b**, **c**) in D1R-SPNs following D1R agonist and/or antagonist injections or those in D2R-SPNs following injection of D2R agonists and/or antagonists, plotted as mean ± s.e.m. across mice. $^*P < 0.05$, $^{***}P < 0.001$, Bonferroni-corrected post hoc comparisons for one-way analysis of variance (ANOVA). Exact P values for all figures are provided in Supplementary Table 1.

activity (a change of about 160 ps) following the administration of the D1R agonist SKF81297 hydrobromide, consistent with D1R-mediated activation of PKA (Fig. 1b, replotted from a previous publication[23]). This D1R-agonist-induced change was not observed in D1R-SPNs that express FLIM-AKAR with a mutated PKA phosphorylation site (FLIM-AKAR(T391A)). By contrast, the D1R antagonist SKF83566 hydrobromide slightly—but significantly (summary statistics in Supplementary Table 1)—increased (by about 17 ps) the FLIM-AKAR lifetime (Fig. 1b), which indicates a reduction in the PKA activity of the D1R-SPNs. Pre-administration of the D1R antagonist largely blocked the D1R-agonist response (Fig. 1b, d), confirming the specificity of the agonist.

In *Adora2a-cre* mice, the D2R agonist sumanirole maleate increased (by about 22 ps) the FLIM-AKAR lifetime, which indicates suppression of the net PKA activity of D2R-SPNs (Fig. 1c). Conversely, the D2R antagonist eticlopride hydrochloride decreased the lifetime by about 16 ps, demonstrating an increase in the net PKA activity of D2R-SPNs (Fig. 1c). Furthermore, pre-injection of a D2R antagonist blocked the effect of the D2R agonist (Fig. 1d). There was a slight change in the lifetime of the phosphorylation-site-mutant AKAR in both D1R-antagonist and D2R-agonist experiments (Fig. 1b, c), which potentially arose from haemodynamic changes induced by the injection.

Changes in PKA in SPNs in response to pharmacological manipulation of DA receptors revealed bidirectional DA-receptor-dependent regulation of PKA activity in SPNs in vivo. Furthermore, the ability of FLiP to detect behaviourally induced changes in the PKA activity of SPNs is suggested by the small (about 10 ps) reduction in FLIM-AKAR lifetime observed in D1R-SPNs at injection, which was independent of the drug administered (Fig. 1b).

Plasticity of DA signals across learning

Because the fluorescence lifetime changes of FLIM-AKAR in response to reward can last 40–60 s[23], we designed a slow-timescale, food-reward reinforced task (Fig. 2a) to investigate how DAN activity, levels of DA in the NAc and PKA activity in SPNs are modulated during learning. We habituated mice to the arena for 1 day and then trained them on the full task for 11 days (days 1–11). On day 12, the reward was omitted from 25% of successful trials to collect 'reward-omission' trials. On day 13, the LED was turned off for the session to collect 'LED-omission' trials, in which mice occasionally managed to perform correct movements despite the absence of the LED cue and received 'unexpected' rewards. The mice learned three key components of the task: (1) staying in the trigger zone to initiate a new trial; (2) running towards the receptacle zone after the LED cue; and (3) waiting in the receptacle zone once having entered it (Extended Data Fig. 2a).

Given that DA release in the NAc can be dissociated from increases in VTA DAN spiking[26], we simultaneously monitored somatic DAN activity in the VTA, the activity of VTA DAN axons in the NAc and the ensuing DA transients in the NAc to compare their patterns across training within the same mice. To simultaneously monitor both DAN activity and DA release in the NAc, we expressed jRCaMP1b[27] in VTA DANs and dLight[28] in NAc (Fig. 2b).

In beginner mice, DA levels in the NAc increased robustly after delivery of the reward but only minimally at the time of the LED cue (Fig. 2c, e, f). In expert mice, the magnitude of DA release after reward was lower than in beginner mice, whereas that after the LED cue was larger (Fig. 2c, e, f). LED-evoked DA release occurred in unrewarded failure trials and also increased across training, and there was no significant difference in LED-evoked DA release in success versus failure trials across training (data not shown). In LED-omission trials, there was no significant increase in DA level at the time of an omitted LED cue ('LO-LED' in Fig. 2d, g); this indicates that the LED-evoked DA response requires the cue, which probably drives reward expectation owing to its learned association with reward. Furthermore, as evidenced by intermediate performance states, the shift in DA release from reward to cue occurred gradually across training (Extended Data Fig. 2i) and correlated with success rate (Fig. 2e, f).

We tested whether positive and negative modulation of DA release occurs in the trained states by examining reward-omission (day 12) and LED-omission (day 13) trials. Consistent with previous studies, DA levels dipped below baseline at the time of expected reward delivery when the reward was omitted ('RO-dip' in Fig. 2d, g). By contrast, the DA peak after reward was larger in rewarded LED-omission trials than in regular rewarded trials (Fig. 2d,g). The dip in DA after reward omission and the larger DA response in unexpectedly rewarded LED-omission trials are consistent with bidirectional modulation of DA by reward expectation.

The patterns of activity in DAN soma and terminals were similar to those of DA levels both across learning and within the trained state (Fig. 2h, Extended Data Fig. 2), and the contributions of behavioural events to each signal were not statistically different in generalized linear models relating the two (Extended Data Fig. 3). In addition, we calculated the correlations among deconvolved fluorescence signals (to account for different kinetics of the signals) or signal peaks, and found that substantial (50–60%) variance in DAN terminal activity and DA release in the NAc are explained by DAN soma activity (Fig. 2i). To test the causal relationships between DAN soma activity and terminal activity, we optogenetically manipulated the former during the behaviour while monitoring the latter. Bidirectional modulation of soma activity blunted the terminal activity response to LED, reward and reward omission (Extended Data Fig. 4).

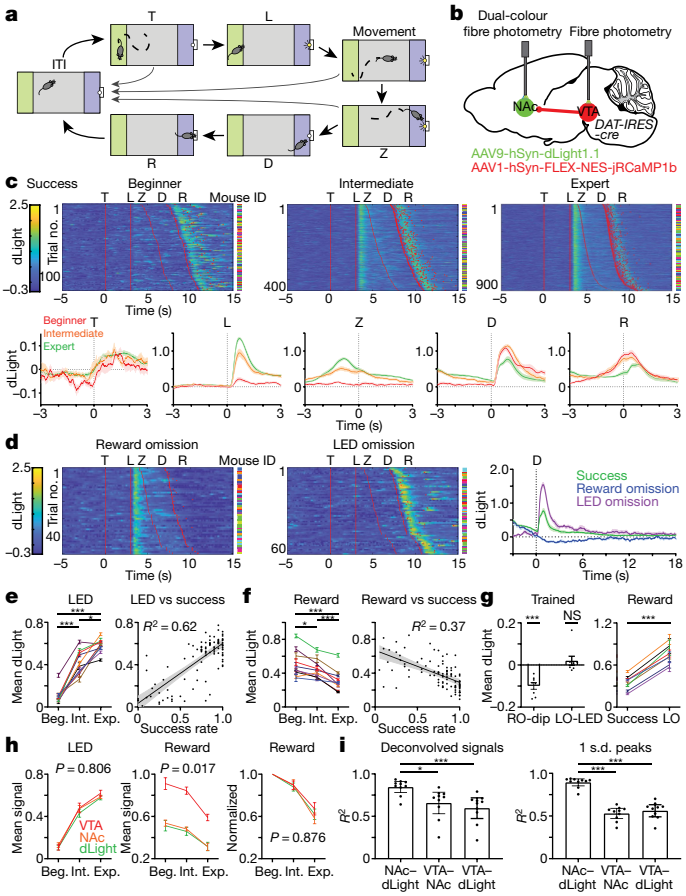


Fig. 2 | Plasticity of DA release and DAN activity dynamics across learning. **a**, After the intertrial interval (ITI) (≥120 s), a mouse can initiate a trial by entering (T) and staying in the trigger zone (green) for ≥3 s, which activates the LED (L). Within 5 s, the mouse must subsequently enter (Z) and stay for ≥3 s in the receptacle zone (blue) to trigger dispensing of the food reward (D), after which the mouse enters the receptacle (R) to collect the pellet. The LED turns off at end of the trial whether due to success (food delivery) or failure (return to the ITI). **b**, Strategy for simultaneous measurement of DA levels and DAN activity. **c**, Top, heat maps of dLight responses in success trials across learning (each row shows a trial). Red lines or dots indicate behavioural timestamps. The mouse identifier (ID) for each row is represented by a different colour. Bottom, average responses across mice aligned to behavioural timestamps, plotted as mean ± s.e.m. across $n = 10$ mice. **d**, dLight signals as in **c** in reward-omission trials (left) and rewarded LED-omission trials (middle) for expert mice. Right, average signals. Plotted as in **c** ($n = 8$ mice). **e**, LED-evoked dLight signals on average for individual mice (plotted as mean ± s.e.m. across trials) at different learning stages (left) and for daily sessions plotted versus success rate (plotted as linear regression fit ± 95% confidence interval) (right). $n = 9$ mice. Beg., beginner; int., intermediate; exp., expert. **f**, Reward-evoked dLight signals plotted as in **e**. $n = 9$ mice. **g**, Responses of individual expert mice to reward omission (RO-dip) and LED omission (LO-LED) (plotted as mean ± s.e.m. across mice) (left), and to reward in regular-success and LED-omission (LO) trials (plotted as in **e**). $n = 8$ mice. NS, not significant. **h**, Average signals to LED (left), reward (middle) and reward normalized to beginner response (right), plotted as mean ± s.e.m. across mice ($n = 9$) and compared by two-way repeated-measures ANOVA for period × signal interaction. **i**, Correlation between pairs of fluorescence channels using deconvolved signals (left) or signal peaks (right) during trials plotted as mean ± 95% confidence interval across $n = 10$ mice. $^*P < 0.05$, $^{***}P < 0.001$ for Bonferroni-corrected post hoc comparisons for one-way repeated-measures ANOVA, one-sample *t*-tests and paired *t*-tests. All *t*-tests are two-sided.

In summary, the task induced both positive and negative modulation of DAN activity and DA release by reward expectation across training stages. These bidirectional changes in DA signals during the slow-timescale behaviour enable examination of how PKA activity in SPNs responds to evolving DA signalling during learning. We performed additional photometry controls, such as recordings in mice expressing enhanced green fluorescent protein (eGFP) or a DA-binding-site-mutant dLight, to examine for potential movement and haemodynamic artefacts, and checked for optical crosstalk between red and green channels, and photobleaching across days (Extended Data Fig. 5).

DA asynchronously modulates SPN PKA

Patterns of DA release in the two hemispheres were similar in this nonlateralized behaviour (Extended Data Fig. 6a, b). Therefore, we performed simultaneous measurements of DA levels and net PKA activity in SPNs by expressing two green fluorescent sensors (dLight and FLIM-AKAR) in different hemispheres (Fig. 3a). Consistent with the above results, dLight responses evoked by the cue and by reward increased and decreased, respectively, across training (Fig. 3b, c). Similarly, net PKA activity in D1R-SPNs increased at the time of reward consumption (red aligned to R in Fig. 3b right) in beginner mice, consistent with the activation of PKA by $G_{\alpha s}$-coupled D1Rs. At expert stages, the increase in the PKA activity of D1R-SPNs shifted to LED onset (Fig. 3b, c). For rewarded trials in trained mice, net PKA activity detected by the sensor probably increased owing to both cue- and reward-evoked DA release (Fig. 3b) as (1) the LED cue increased PKA activity in unrewarded failure trials and in reward-omission trials; and (2) the magnitude of net PKA activation was larger in rewarded compared to unrewarded trials. Furthermore, PKA activation in D1R-SPNs was larger in rewarded LED-omission trials of expert mice than in regular rewarded trials, consistent with the larger DA release evoked by an unexpected reward triggering greater PKA activation (Fig. 3d, e).

By contrast, net PKA activity in D2R-SPNs was not strongly modulated in beginner mice (Fig. 3b). Furthermore, although a slight reduction in net PKA activity in D2R-SPNs after the LED cue is consistent with PKA inhibition by $G_{\alpha i/o}$-coupled D2Rs activated by DA, this pattern did not change across learning (Fig. 3b, c). Therefore, this dip in PKA activity was probably not caused by DA, as the release pattern of DA changed markedly across learning. By contrast, failure trials of intermediate and expert mice—in which DA levels significantly dropped below baseline—significantly increased net PKA activity in D2R-SPNs (Fig. 3b, c). A similar pattern of FLIM-AKAR lifetime in D2R-SPNs occurred in reward-omission trials, consistent with this signal reflecting activation of PKA by DA decreasing below baseline (Fig. 3d, e). Furthermore, the modulations of FLIM-AKAR lifetime observed during behaviour occurred in individual trials (Extended Data Fig. 7) and did not reflect movement artefacts (Extended Data Fig. 6c). Thus, PKA activities in D1R- and D2R-SPNs respond to different DA dynamics and are not strongly modulated at the same time: PKA in D1R-SPNs is activated at early learning stages by rewards as well as by each reward-predictive cue and reward after learning (which increase DA levels), whereas PKA in D2R-SPNs is activated only at late learning stages by failures to achieve expected rewards (which decrease DA below baseline).

With the exception of the small reduction in net PKA activity in D2R-SPNs in success trials, the net PKA activity patterns in SPNs could be explained by the evolution of DA dynamics across learning. To test causality in this relationship (Fig. 3f, g), we exploited the DA-receptor-targeting pharmacology shown in Fig. 1. Both LED- and reward-driven increases of net PKA activity in D1R-SPNs were largely blocked by the D1R antagonist SKF83566 hydrobromide. This demonstrates that PKA activation in D1R-SPNs that is correlated with DA release is indeed dependent on this DA receptor. Similarly, the activation of PKA in D2R-SPNs by reward omission was blocked by the D2R antagonist eticlopride hydrochloride, which indicates that this PKA activation is mediated by D2Rs and requires basal DA binding to D2Rs. By contrast, the small reduction in net PKA activity in D2R-SPNs after reward consumption was not blocked by the D2R antagonist.

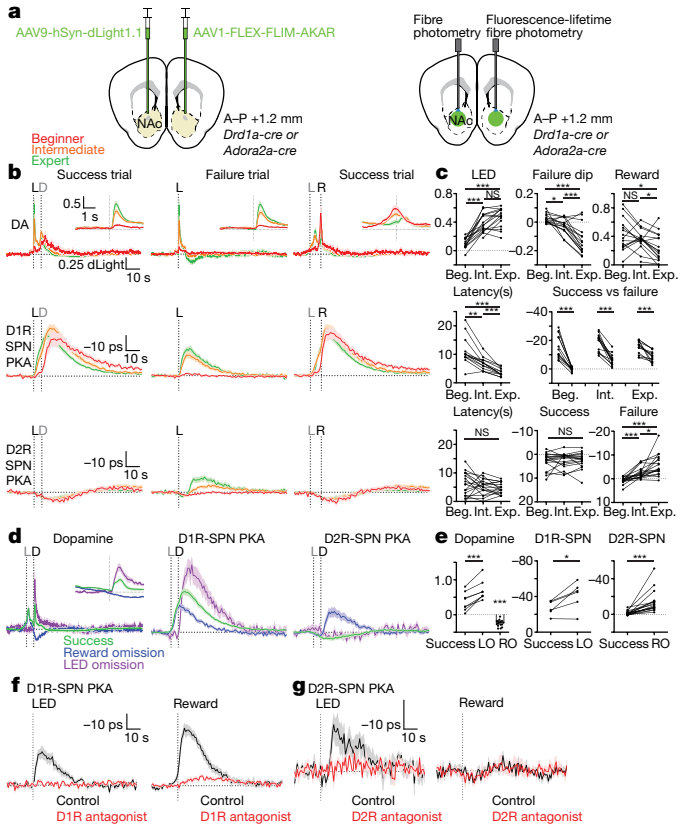


Fig. 3 | PKA activities in D1R- and D2R-SPNs are dynamically modulated and follow patterns of DA during learning. a, Strategy to measure PKA activity of SPNs and DA release. **b**, Average responses of dLight of *Drd1a-cre* and *Adora2a-cre* mice ($n = 16$) (top), FLIM-AKAR in D1R-SPNs in *Drd1a-cre* mice ($n = 14$) (middle) and FLIM-AKAR in D2R-SPNs in *Adora2a-cre* mice ($n = 18$) (bottom). Alignments are to (dashed vertical line) LED onset (L) (black) for success (left) and failure (middle) trials or receptacle entry (R) (black) for success trials (right). The average times across mice of pellet dispensing (D) and LED onset (L) are in grey. **c**, Top, dLight signal amplitudes after LED onset (left), in failure trials at typical time of reward delivery (middle) and to reward (right). Middle, D1R-SPN PKA response latency (to 50% peak) and mean amplitude for success versus failure trials. Bottom, PKA response latency (to 50% valley) and mean amplitudes for success and failure trials, in D2R-SPNs. **d**, dLight and FLIM-AKAR responses in reward-omission and rewarded LED-omission trials aligned to pellet dispensing (D) (black). Average time of LED onset (L) is in grey. **e**, Average responses to rewards in regular-success and rewarded LED-omission (LO) trials, or at time of expected reward delivery in reward-omission (RO) trials for dLight (left), PKA in D1R-SPNs (middle) and PKA D2R-SPNs (right). **f**, Effects of D1R antagonism on PKA activity transients ($n = 4$) in D1R-SPNs, aligned to onset of an LED cue previously associated with food delivery but with no food delivered (left) (five trials per mouse) or to a free (that is, no action required) reward aligned to receptacle entry (right) (ten trials per mouse). **g**, As in **f**, for FLIM-AKAR in D2R-SPNs ($n = 4$ mice) and D2R antagonism. All graphs are plotted as mean ± s.e.m. across mice, except for **c** and **e** in which dots are mouse averages; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ for Bonferroni-corrected post hoc comparisons for one-way repeated-measures ANOVA, one sample *t*-tests and paired *t*-tests. All *t*-tests are two-sided. The numbers of trials and mice per condition are given in Extended Data Fig. 7.

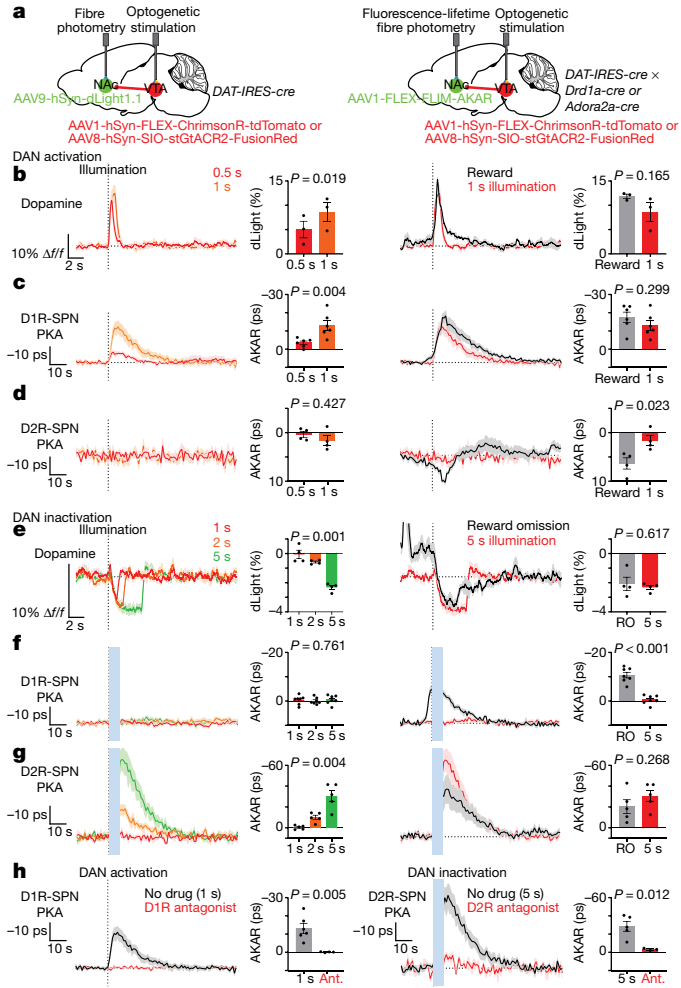


Fig. 4 | Transient changes in DAN activity are sufficient to modulate PKA activity in SPNs. a, Strategy for modulating DAN activity and measuring DA release (left) or PKA activity in SPNs (right). **b**, dLight responses of *DAT-IRES-cre* mice ($n = 3$) to DAN activation (20 Hz) for different durations (left) and of naive animals to food reward compared to 1 s illumination (right). Dashed line, time of illumination onset or reward delivery. Fluorescence changes are plotting the change from baseline divided by baseline ($\Delta f/f$). **c**, As in **b**, showing FLIM-AKAR responses in D1R-SPNs of *DAT-IRES-cre;Drd1a-cre* mice ($n = 6$). **d**, As in **b**, showing FLIM-AKAR responses in D2R-SPNs of *DAT-IRES-cre;Adora2a-cre* mice ($n = 4$). **e**, dLight responses of *DAT-IRES-cre* mice ($n = 4$) to DAN inactivation (stGtACR2 continuous illumination for different durations) (left) and of trained mice to reward omission compared to 5 s inactivation (right). Dashed line, time of illumination onset or expected reward delivery. Baseline dLight signal in reward omission is due to the response to the antecedent LED cue. **f**, As in **e**, for FLIM-AKAR responses in D1R-SPNs of *DAT-IRES-cre;Drd1a-cre* mice ($n = 7$). **g**, As in **e**, for FLIM-AKAR responses in D2R-SPNs of *DAT-IRES-cre;Adora2a-cre* mice ($n = 5$). **h**, Responses of FLIM-AKAR responses in D1R-SPNs to DAN activation (1 s) without ($n = 6$) or with ($n = 4$) injection of D1R antagonist (left), and in D2R-SPNs to DAN inactivation (5 s) without ($n = 5$) and with ($n = 3$) injection of D2R antagonist (right). Ant., antagonist. All graphs are plotted as mean ± s.e.m. across mice. Dot, average for each mouse. Paired *t*-test, unpaired *t*-test and one-way repeated-measures ANOVA. All *t*-tests are two-sided. Blue bars, laser illumination.

To further test the sufficiency in the relationship between transients of DA and PKA activity in SPNs, we investigated the effect of optogenetically activating (ChrimsonR[29]) or inactivating (stGtACR2[30]) DANs on PKA activity in SPNs (Fig. 4a). We confirmed the ability of the two opsins to bidirectionally modulate DA levels in a graded manner (Fig. 4b, e). An activation protocol that achieved peak DA release, similar to a food-reward response, increased net PKA activity in D1R-SPNs in a D1R-dependent manner to a similar extent as the reward (Fig. 4b, c, h). Similar effects were achieved by stimulating DAN axons in the NAc and by ramping DA to mimic patterns observed during goal-directed movements[31] (Extended Data Fig. 8b, c). Furthermore, the relationship between PKA activation and DA release was nonlinear (Extended Data Fig. 8d), which suggests that PKA activation depends on both the amount and the kinetics of DA release.

By contrast, net PKA activity in D2R-SPNs was minimally modulated by DAN activation (Fig. 4d), which provides support for the hypothesis that—consistent with its insensitivity to D2R antagonism (Fig. 3g)—the small reduction in PKA activity in D2R-SPNs that is induced by reward is independent of DA. Furthermore, the insensitivity of PKA in D2R-SPNs to DAN activation and increasing DA suggests that D2R-dependent inhibition of PKA is close to maximal at basal DA. Non-physiological levels of DA did modulate net PKA activity in D2R-SPNs (Extended Data Fig. 8a).

DAN inactivation had converse effects. Net PKA activity in D1R-SPNs was not significantly changed in response to DAN inactivation (Fig. 4f); this indicates a minimal engagement of D1Rs by basal DA, which makes D1Rs unresponsive to DA dips. Importantly, optogenetic inhibition of DANs isolated the effect of DA decreases on PKA in D1R-SPNs—unlike in reward-omission trials, in which the cue-evoked DA increase precedes the dip. Conversely, net PKA activity in D2R-SPNs significantly increased in a D2R-dependent manner during DAN inactivation that decreases DA to a similar extent as reward omission in trained mice (Fig. 4e,g,h), consistent with a basal engagement of D2Rs by DA that allows unbinding of DA from D2Rs to disinhibit PKA in D2R-SPNs.

There were two results from the optogenetic experiments that were not consistent with the pharmacological manipulations. First, the inactivation of DANs—which dipped DA below baseline to a similar extent as reward omission—did not affect PKA activity in D1R-SPNs (Fig. 4f), but D1R antagonism suppressed PKA in D1R-SPNs (Fig. 1b). Second, activation of DANs to increase DA levels to a similar extent as food rewards did not affect PKA activity in D2R-SPNs (Fig. 4d), but D2R agonism reduced PKA activity in D2R-SPNs (Fig. 1c). Besides possible nonspecific and circuit-wide drug effects, these discrepancies might result from stronger and longer-lasting effects of the D1R antagonist and D2R agonist as compared to physiologically induced short-term changes in DA.

PKA inhibition in SPNs slows learning

We investigated whether selective inhibition of PKA activity in each neuron class affects learning by virally overexpressing PKI_{α} (hereafter, PKI), an endogenous inhibitory peptide that blocks PKA activation in vitro[24] and in vivo[23]. PKI expression in D1R-SPNs altered learning, as evident from daily averages of time in trigger zone, speed after LED and entering failure rate (Extended Data Fig. 9b–d). The effect of PKA inhibition in D1R-SPNs appeared in early training (days 1–3), consistent with PKA activation in D1R-SPN by reward in beginner mice. Additional differences in trigger-zone time after day 3 suggest that PKA activation in D1R-SPNs continues to contribute to learning in intermediate mice (Extended Data Fig. 9b). PKA inhibition in D1R-SPNs also reduced overall speed (Extended Data Fig. 9i). However, this effect cannot fully explain the above effects because the average speed after LED onset was reduced for the PKI group in the first half of training only (Extended Data Fig. 9c), which indicates that mice were physically capable of fulfilling task criteria.

The effects of PKA inhibition in D2R-SPNs were significant at later stages: speed after LED and entering failure rate in days 4–7 were altered (Extended Data Fig. 9f, g). This effect is consistent with PKA activation in D2R-SPNs by failure to receive rewards in intermediate and expert mice (Fig. 3b). On the other hand, there was no significant difference in time in trigger zone (Extended Data Fig. 9e), consistent with the lack of DA modulation by failure to initiate LED (LED omission). Although PKA in D2R-SPNs was strongly activated during reward-omission trials, PKA inhibition in D2R-SPNs did not alter extinction of task-related behaviours (data not shown). This is consistent with a recent demonstration that extinction involves neither DA dips nor D2R-SPNs in the NAc[20], and instead depends on different parts of the striatum[32].

Cell-type-specific PKA inhibition produced mild effects on behaviour, and not a complete impairment of learning. This may be because we designed a focal manipulation (that is, inhibition of PKA in specific cells in a subregion of the NAc). In addition, the subtlety of the effects could also be due to (1) the incomplete blockade of PKA in PKI-expressing cells; (2) the engagement of PKA-independent cellular plasticity in SPNs[33]; and (3) the contribution of other brain regions that act in parallel to the NAc or compensate for the loss of PKA signalling in the NAc[34].

Discussion

Despite the results of pioneering work[35–37], directly measuring the downstream effects of positive and negative DA transients on SPNs has been difficult to achieve owing to the challenges of monitoring intracellular signalling in behaving animals. By using FLiP, here we monitored net PKA activity in SPNs during behaviour with a temporal resolution of 1 s and have shown for the first time—to our knowledge—that PKA activity in D1R- and D2R-SPNs dynamically follows the patterns of DA release evoked by conditioned and unconditioned events. Furthermore, we have established the causality of the relationship between DA and PKA activity of SPNs by combining pharmacology and optogenetics with FLiP. Despite the potential contributions of many G-protein-coupled receptors and Ca^{2+}-dependent adenylyl cyclases, DA—at times of its phasic modulation—largely dictates the state of PKA in SPNs. Furthermore, our findings provide in vivo evidence for a mechanistic model of the modulation of PKA in SPNs by DA (Extended Data Fig. 10), which explains the selective sensitivity of PKA activity in D1R- and D2R-SPNs to increases and decreases in DA, respectively (owing to the different basal occupancy of the receptors). Our demonstration of response of PKA in D2R-SPNs to DA dips during behaviour is especially valuable in light of the recent finding that DA dips induce PKA-dependent spine growth in D2R-SPNs[20].

As a consequence of their temporally dissociated DA-dependent modulation of PKA, SPNs of the direct (striatomesencephalic) and indirect (striatopallidal) pathways may have independent functions during learning. On the basis of our results, we endorse that the former promotes an initial association between an action and an outcome and the latter refines the learned behaviour once the association is established. This model is consistent with previous theoretical[38,39] and conceptual[1,40] models, as well as with recent experimental results[20] that support the same core concept of opponent reinforcement learning.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-03050-5.

1. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* **68**, 815–834 (2010).
2. Kravitz, A. V. & Kreitzer, A. C. Striatal mechanisms underlying movement, reinforcement, and punishment. *Physiology (Bethesda)* **27**, 167–177 (2012).
3. Vidal-Gadea, A. G. & Pierce-Shimomura, J. T. Conserved role of dopamine in the modulation of behavior. *Commun. Integr. Biol.* **5**, 440–447 (2012).
4. Steinberg, E. E. et al. Positive reinforcement mediated by midbrain dopamine neurons requires D1 and D2 receptor activation in the nucleus accumbens. *PLoS ONE* **9**, e94771 (2014).
5. Hikida, T., Kimura, K., Wada, N., Funabiki, K. & Nakanishi, S. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* **66**, 896–907 (2010).
6. Tsai, H. C. et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080–1084 (2009).
7. Steinberg, E. E. et al. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966–973 (2013).
8. Saunders, B. T., Richard, J. M., Margolis, E. B. & Janak, P. H. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat. Neurosci.* **21**, 1072–1083 (2018).
9. Coddington, L. T. & Dudman, J. T. The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nat. Neurosci.* **21**, 1563–1573 (2018).
10. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
11. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).

12. Eshel, N., Tian, J., Bukwich, M. & Uchida, N. Dopamine neurons share common response function for reward prediction error. *Nat. Neurosci.* **19**, 479–486 (2016).
13. Day, J. J., Roitman, M. F., Wightman, R. M. & Carelli, R. M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020–1028 (2007).
14. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
15. Gerfen, C. R. et al. D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250**, 1429–1432 (1990).
16. Kupchik, Y. M. et al. Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. *Nat. Neurosci.* **18**, 1230–1232 (2015).
17. Skeberdis, V. A. et al. Protein kinase A regulates calcium permeability of NMDA receptors. *Nat. Neurosci.* **9**, 501–510 (2006).
18. Lee, H. K. et al. Phosphorylation of the AMPA receptor GluR1 subunit is required for synaptic plasticity and retention of spatial memory. *Cell* **112**, 631–643 (2003).
19. Yagishita, S. et al. A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* **345**, 1616–1620 (2014).
20. Iino, Y. et al. Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature* **579**, 555–560 (2020).
21. Lau, G. C., Saha, S., Faris, R. & Russek, S. J. Up-regulation of NMDAR1 subunit gene expression in cortical neurons via a PKA-dependent pathway. *J. Neurochem.* **88**, 564–575 (2004).
22. Nayak, A., Zastrow, D. J., Lickteig, R., Zahniser, N. R. & Browning, M. D. Maintenance of late-phase LTP is accompanied by PKA-dependent increase in AMPA receptor synthesis. *Nature* **394**, 680–683 (1998).
23. Lee, S. J., Chen, Y., Lodder, B. & Sabatini, B. L. Monitoring behaviorally induced biochemical changes using fluorescence lifetime photometry. *Front. Neurosci.* **13**, 766 (2019).
24. Chen, Y., Saulnier, J. L., Yellen, G. & Sabatini, B. L. A PKA activity sensor for quantitative analysis of endogenous GPCR signaling via 2-photon FRET-FLIM imaging. *Front. Pharmacol.* **5**, 56 (2014).
25. Chen, Y. et al. Endogenous Gαq-coupled neuromodulator receptors activate protein kinase A. *Neuron* **96**, 1070–1083.e5 (2017).
26. Mohebi, A. et al. Dissociable dopamine dynamics for learning and motivation. *Nature* **570**, 65–70 (2019).
27. Dana, H. et al. Sensitive red protein calcium indicators for imaging neural activity. *eLife* **5**, e12727 (2016).
28. Patriarchi, T. et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360**, eaat4422 (2018).
29. Klapoetke, N. C. et al. Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014).
30. Mahn, M. et al. High-efficiency optogenetic silencing with soma-targeted anion-conducting channelrhodopsins. *Nat. Commun.* **9**, 4125 (2018).
31. Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M. & Graybiel, A. M. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**, 575–579 (2013).
32. Matamales, M. et al. Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. *Science* **367**, 549–555 (2020).
33. Jiang, S. Z. et al. NCS-Rapgef2, the protein product of the neuronal *Rapgef2* gene, is a specific activator of D1 dopamine receptor-dependent ERK phosphorylation in mouse brain. *eNeuro* **4**, ENEURO.0248-17.2017 (2017).
34. Ilango, A. et al. Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *J. Neurosci.* **34**, 817–822 (2014).
35. Goto, A. et al. Circuit-dependent striatal PKA and ERK signaling underlies rapid behavioral shift in mating reaction of male mice. *Proc. Natl Acad. Sci. USA* **112**, 6718–6723 (2015).
36. Yamaguchi, T. et al. Role of PKA signaling in D2 receptor-expressing neurons in the core of the nucleus accumbens in aversive learning. *Proc. Natl Acad. Sci. USA* **112**, 11383–11388 (2015).
37. Ma, L. et al. A highly sensitive A-kinase activity reporter for imaging neuromodulatory events in awake mice. *Neuron* **99**, 665–679.e5 (2018).
38. Collins, A. G. E. & Frank, M. J. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol. Rev.* **121**, 337–366 (2014).
39. Gurney, K. N., Humphries, M. D. & Redgrave, P. A new framework for cortico-striatal plasticity: behavioural theory meets in vitro data at the reinforcement-action interface. *PLoS Biol.* **13**, e1002034 (2015).
40. Gerfen, C. R. & Surmeier, D. J. Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34**, 441–466 (2011).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Mice

Experimental manipulations were performed in accordance with protocols approved by the Harvard Standing Committee on Animal Care, following guidelines described in the US National Institutes of Health Guide for the Care and Use of Laboratory Animals. *Drd1a-cre* (B6.FVB(Cg)-Tg(Drd1-cre)EY262Gsat/Mmucd, 030989-UCD) and *Adora2a-cre* (B6.FVB(Cg)-Tg(Adora2a-cre)KG139Gsat/Mmucd, 036158-UCD) mice on C57BL/6J backgrounds[41] were acquired from MMRRC UC Davis. *DAT-IRES-cre* (B6.SJL-Slc6a3tm1.1(cre)Bkmn/J, 006660)[42] and C57BL/6J (000664) mice were acquired from the Jackson Laboratory. *DAT-IRES-cre;Drd1a-cre* and *DAT-IRES-cre;Adora2a-cre* mice were bred in-house by crossing heterozygous parent lines. All transgenic mice used for experiments were heterozygous for the relevant *cre* allele. Mouse ages were 2–4 months, and male and female mice were used in approximately equal proportion. Mice were housed on a 12 h/12 h dark/light reversed cycle. No sample size precalculation was performed. For PKI experiments, we randomly assigned mice to PKI and GFP group in a manner such that each of a pair matched in age and sex was randomly assigned to a different group. For PKI experiments, no blinding was attempted.

Viruses

Recombinant adeno-associated viruses (AAVs of serotype 1, 8 and 9) were used to express transgenes of interest in either Cre-recombinase-dependent or -independent manner. AAVs were packaged by commercial vector core facilities (Addgene, Boston Children's Hospital Vector Core, Janelia Vector Core, Penn Vector Core and UNC Vector Core) and stored at −80 °C upon arrival. Viruses were used at a working concentration of 10^{12} to 10^{14} genomic copies per ml. Three hundred nl of virus was used for all experiments, except for PKI experiments for which we used 600 nl of virus at 1.2×10^{14} genomic copies per ml and 300 nl of virus at 4×10^{13} genomic copies per ml for *Drd1a-cre* and *Adora2a-cre*, respectively (we used a lower volume and titre for the latter owing to greater toxicity in this line). Most viral plasmids are available in Addgene: AAV-FLEX-FLIM-AKAR (no. 60445), AAV-FLEX-FLIM-AKART391A (no. 60446), AAV-CAG-dLight1.1 (no. 111067, hSyn construct available upon request to L. Tian), AAV-hSyn-FLEX-NES-jRCaMP1b (no. 100850), AAV-hSyn-FLEX-ChrimsonR-tdTomato (no. 62723), AAV-hSyn-SIO-stGtACR2-FusionRed (no. 105677), AAV-FLEX-PKIalpha-IRES-nls-mRuby2 (no. 63059), AAV-Cag-FLEX-eGFP (no. 59331) and AAV-Syn-Flex-GCaMP6f (no. 100833). AAV-hSyn-dLight^{D103A} is available upon request to L. Tian.

Surgery

Inhaled isoflurane was used in anaesthesia. Virus was stereotactically injected into either the NAc core (anteroposterior +1.2 mm, mediolateral ±1.3 mm relative to bregma; dorsoventral 4.1 mm below brain surface) or VTA (anteroposterior −3.3 mm, mediolateral +0.48 mm, dorsoventral 4.5 mm). For fibre photometry or optogenetic experiments, an optical fibre (MFC_200/230-0.37_4.5mm_MF1.25_FLT mono fibre optic cannula, Doric Lenses) was implanted 200 μm above the injection site.

Behaviour

To motivate mice to perform the visual-cue-guided task, we food-restricted mice such that they remained at about 80–90% of their initial weight. Mice were given about 2–3 g of regular chow daily, in addition to the variable number of 20 mg dustless precision chocolate flavour pellets (F05301, Bio Serv) consumed during the task. Food restriction was started at least one day before commencing the behavioural training. During the task, a mouse (sometimes connected to a patch cord) was allowed to freely move inside an 8 × 16 inch box, which contained a pellet receptacle and a white LED on one of the short walls. Receptacle entry was detected by an infrared sensor installed inside the receptacle. Mouse movements were captured by cameras (WV-CP504, Panasonic or FL3-U3-13E4M, PointGrey) connected to Ethovision 11.5 or Bonsai 2.3 software that controlled all behavioural apparatuses and were synchronized with MATLAB 2012a software used to acquire photometry data.

The behavioural task structure was as follows. After the enforced 120-s ITI, the mouse can self-initiate a trial by entering the trigger zone (a zone opposite to the wall containing the pellet receptacle, indicated in green in Fig. 2a but not marked in the actual behaviour box) and staying in the zone for 3 consecutive seconds. If the mouse enters the trigger zone during the enforced ITI or exits the trigger zone before 3 s after entering the zone, the trial is not initiated. If the mouse succeeds in initiating the trial, the LED above the receptacle turns on, signalling the start of the trial. Once the LED turns on, the mouse must enter the LED zone (a zone near the wall containing the pellet receptacle, indicated in blue in Fig. 2a but not marked in the box) within 5 s. If the mouse fails to enter the LED zone within 5 s, the trial is terminated and the LED turns off. If the mouse enters the LED zone in time, it has to stay in the LED zone for additional 3 consecutive seconds. After this enforced waiting period, a single 20 mg pellet is dispensed, and the LED turns off. If the mouse prematurely exits the LED zone during the waiting period, the trial is terminated and the LED turns off. After termination of the trial either by a success or a failure, the next enforced ITI of 120 s starts. Mice were trained for 11 days starting after 1 day of a 40-min habituation session during which they were given 10 free pellets in the receptacle. On day 12, the reward was omitted from 25% of successful trials to collect 'reward-omission' data. On day 13, the LED was turned off for the entire session to collect 'LED-omission' data. Most mice reached 90% success rate (number of success trials/number of total trials) in 9 days (Extended Data Fig. 2a).

Each behaviour session was run until a mouse initiated 20 trials (including success and failure trials) or a 2-h time limit was reached, except on reward-omission days (day 12) when the session continued until 5 reward-omission trials were collected. Because mice had to self-initiate trials, session durations were variable (50 min to 2 h). Most sessions included 20 trials except the first few sessions at the start of training and the extinction sessions of PKI experiments (9 out of 806 sessions for photometry experiments, 15 out of 611 sessions for PKI experiments).

To compare signals across different stages of learning across all mice, we defined a beginner stage as days in which the success rate was below 50% of the maximum (plateau) success rate, an intermediate stage as days in which the success rate was between 50 and 90% of the maximum success rate, and an expert stage as days in which the success rate was above 90% of the maximum success rate. A highly sensitive optical system allowed us to use dim excitation (about 5–16 μW) and perform fibre photometry on all days to track progressive changes in neural signals.

FLiP

FLiP was carried out with the optical system described in Extended Data Fig. 1 and using the previously described method[23]. In brief, all filters in the system were purchased from Semrock. A pulsed laser (BDS-473-SM-FBE Becker and Hickl (BH) operating at 50 MHz) was used as the light source to excite FLIM-AKAR. For fluorescence detection, a high-speed hybrid photomultiplier tube (PMT, HPM-100-07-Cooled, BH) controlled by DCC-100-PCI (BH) was used. The hybrid PMT was connected to an SPC-830 (BH), a time-correlated single-photon counting (TCSPC) board, which detects the time delay between the pulsed excitation and the photon detection by the PMT. The data were collected

by custom software written in MATLAB 2012a, which calculated the average lifetime of detected photons at 1-s intervals. This interval for average lifetime measurements was empirically determined to have enough photons to accurately estimate the lifetime (>200,000 photons per measurement) without running into the photon count limit (about 1,000,000 photons per s) of the TCSPC board. The typical excitation power needed to generate the appropriate rate of photons (about 400 kHz) for TCSPC was about 0.6–1 μW (measured at the output end of the patch cord).

To estimate the change in lifetime of FLIM-AKAR, we measured the average lifetime for each 1-s time bin by calculating the mean photon arrival time (the population mean of the delay between the pulsed excitation and the fluorescence photon arrival as described) using the following equation[43]:

$$\tau = t - t_0 = \frac{\int tF(t) \cdot dt}{\int F(t) \cdot dt} - t_0,$$

in which $F(t)$ is the photon count of a fluorescence-lifetime decay curve at time bin t, and t_0 is the offset of the lifetime histogram, which can be estimated by fitting a double exponential curve to a lifetime histogram. We performed this calculation for the 8-ns time range in a lifetime histogram as this interval was minimally contaminated by a secondary fluorescence peak resulting from the autofluorescence of a fibre. The length of the patch cord was chosen to maximize the time separation between the sensor fluorescence peak and the fibre autofluorescence peak (about 10-ns time delay for light to travel from the one end of the patch cord to the other end and back). The lifetime of FLIM-AKAR was reported as a change in lifetime (delta lifetime), which was calculated by subtracting the average lifetime of a baseline period (a period before the event of interest) from the average lifetime transient. Delta lifetime was plotted in 1-s time bins for behavioural experiments and in 10-s averages for pharmacology experiments.

Fibre photometry and optogenetics

Both fibre photometry (fluorescence intensity fibre photometry) and optogenetics were carried out with the optical system described in Extended Data Fig. 1. All filters in the system were purchased from Semrock except for a Doric Minicube (FMC5_E1(465-480)_F1(500-540)_E2(555-570)_F2(580-680)_S, Doric Lenses) that houses multiple dichroic mirrors and filters.

For fibre photometry, 470-nm (M470F3, Thorlabs) and 565-nm (M565F3, Thorlabs) LEDs were used to excite dLight (or GCaMP) and jRCaMP, respectively. Both LEDs were frequency-modulated by a digital signal processor system (RX8-5-12, Tucker-Davis Technology) to carry out locked in amplification of PMT outputs. The average power levels of LEDs (measured at the output end of the patch cord) were 9.3 μW, 5.4 μW, 16.2 μW, 5.6 μW and 14.5 μW for dLight, soma jRCaMP, terminal jRCaMP, soma GCaMP and terminal GCaMP excitation. For fluorescence detection, H7422-40 (Hamamatsu) and H10770(P)A-40 (Hamamatsu) PMTs were used. PMTs were connected to a low-noise current preamplifier (SR570, Stanford Research Systems). The signal generated from SR570 was locked-in-amplified by RX8-5-12 using the frequency of the LED used to excite the sensor that fluoresces in the light spectrum assigned to a corresponding PMT. The locked-in-amplified signal was collected by a data acquisition board (PCI-6115, National Instruments), which was controlled by the custom software written in MATLAB 2012a. The raw fluorescence data was collected at 1 kHz. It was subsequently smoothed by a moving average filter (width of 200 ms) and down-sampled to 100 Hz. The relative change in fluorescence $\Delta f(t)/f_0 = (f(t) - f_0)/f_0$ was calculated using f_0 equal to the average of the baseline period (20 s before the trigger zone entry). For comparisons and averaging across mice, $\Delta f/f$ of an individual mouse was normalized to the 99th percentile of $\Delta f/f$ value across all sessions for the mouse.

For activation of Chrimson, a 593.5-nm laser (SKU: YL-593-00100-CWM-SD-03-LED-F, Optoengine) was used with 2-ms laser pulses delivered at 20 Hz with exception of DA ramping experiments (Extended Data Fig. 8). For DA ramping experiments, we increased the frequency of the light pulses (2-ms width) gradually from 24 Hz to 34 Hz for 3 s, 16 Hz to 34 Hz for 5 s, and 4 Hz to 30 Hz for 7 s by changing the frequency every 500 ms. Reported laser powers were measured by a digital optical power meter (PM100D, Thorlabs) at the end of the patch cable, while the laser was operating in a continuous mode. For activation of stGtACR2, a 473-nm laser (MBL-III-473, Optoengine) was used. For all stGtACR2 activation experiments, the laser was operated in a continuous mode (without pulsing). Reported laser powers were measured in the same manner as in the Chrimson experiments.

Pharmacology

The following concentrations of drugs were used for intraperitoneal injection: D1R agonist (SKF 81297 hydrobromide, 10 mg kg^{-1}), D1R antagonist (SKF 83566 hydrobromide, 3 mg kg^{-1}), D2R agonist (sumanirole maleate, 4 mg kg^{-1}), D2R antagonist (eticlopride hydrochloride, 0.5 mg kg^{-1}). All drugs were dissolved in sterile saline. Drug solutions were intraperitoneally injected in 0.1 ml solution per 10 g of mouse using a fine needle insulin syringe (BD 324911, BD Bioscience). On average, it took about 30–60 s from the beginning of scruffing a mouse to the end of intraperitoneal injection.

Deconvolution of fluorescence intensity photometry data

To gain a better understanding of the underlying neural activity generating the measured fluorescence transients, we deconvolved fluorescence signals into population events using a previously developed one-dimensional constrained deconvolution algorithm[44], which assumes that the sensor fluorescence follows an autoregressive process. Before deconvolution, we smoothed the raw fluorescence data collected at 1 kHz by a moving average filter (200-ms width) and down-sampled the data to 10 Hz. We subsequently calculated $\Delta f/f$ using the bottom 5th percentile value of a rolling time window of 300 s as f_0. This allowed us to perform deconvolution on fluorescence signal for an entire session without introducing a trial structure to the data. The deconvolution of $\Delta f/f$ was performed using a conic programming method and a second order autoregressive process. Correlations between deconvolved signals were calculated by computing a correlation between the number of population events in each 1-s time interval.

Encoding model

To more objectively reveal the relationships between behaviour and fluorescence signals, we created a generalized linear model to predict the deconvolved fluorescence signal (dLight and jRCaMP) from observed stimuli and behavioural parameters. We processed both observed stimuli and behavioural parameters into time bin of 100 ms to match the time bin of the deconvolved fluorescence signal. We constructed additional explanatory variables by introducing multiple time shifts to the observed stimulus and behavioural parameter variables. For continuous variables (speed, acceleration, rotation and position), time shifts of −0.3 to 0.3 s were introduced; for event variables (movement initiation, cue, reward delivery and receptacle entry), time shifts of −0.3 to 1.0 s were introduced. We reorganized the data from a session into a trial structure to have a fixed distribution of the number of trials when we split the data into fitting and testing sets. Model parameters (coefficients) were learned from fitting sets and evaluated using testing sets (fivefold cross validation). We performed Lasso on the data from all sessions of each learning stage and each mouse to find a minimum set of explanatory variables that allows the mean squared error to be within 1 standard error from the minimum mean squared error (a standard error of mean squared error was calculated from five cross-validation sets). We decided on the final common set of explanatory variables by selecting all variables selected

by Lasso across different learning stages, mice and signals. With the final set of explanatory variables, we performed a linear regression. The contribution of each variable category (kinematics, movement initiation, position, cue, reward, accuracy and previous trial) was calculated by setting the coefficients of all the variables assigned to each category as zero and computing the correlation between the actual fluorescence signal and the signal predicted by the model: contribution = $(R^2_{\text{full model}} - R^2_{\text{partial model}})/R^2_{\text{full model}}$.

Histology

Virus-injected mice were euthanized and perfused transcardially with PBS followed by 4% PFA (in PBS). After >24-h post-fix in 4% PFA, brains were sliced at 50-µm thick using a vibrating blade microtome (Leica Biosystems VT1000S), mounted on glass slides with a DAPI mounting medium (Vectashield, H-1200), and imaged under a wide-field microscope with a 10× objective (VS120, Olympus). Images were acquired through OlyVIA 2.9 and processed via ImageJ 1.52i, which was also used for cell counting with local contrast enhancement (CLAHE) and particle analysis.

Statistics

For all data presentation, we first averaged trials for individual mice, and then averaged across mouse averages. All error bars are s.e.m. across mouse averages, unless stated otherwise in the legend. All pairwise comparisons were two-sided. All multiple comparisons were Bonferroni-corrected. Exact *P* values are reported in Supplementary Table 1. R^2 represents a Pearson's correlation coefficient. For analysis of a trend of a single group across multiple time periods, one-way repeated-measures ANOVA was used to take into account of repeated measurements from the same subjects, (if needed) followed by Bonferroni post hoc comparisons. Similarly, for analysis of a trend of multiple groups across multiple time periods, two-way repeated-measures ANOVA was used, followed by Bonferroni post hoc comparisons. For the analyses in dLight-jRCaMP (Fig. 2), dLight of dLight-AKAR (Fig. 3) and D1R-SPN AKAR of dLight-AKAR (Fig. 3) experiments, 1 out of 10, 1 out of 16, and 1 out of 14 subjects, respectively, were dropped from the repeated-measures ANOVA owing to a missing value in serial measurements (no measurement made). In the case of datasets with multiple missing values, we analysed the data instead by fitting a mixed model (one-way) as implemented in GraphPad Prism 8.0[45] (Extended Data Fig. 2i). This mixed model uses a compound symmetry covariance matrix and is fit using restricted maximum likelihood. In the absence of missing values, this method gives the same *P* values and multiple comparisons tests as repeated-measures ANOVA. In the presence of missing values (missing completely at random), the results can be interpreted as with a repeated-measures ANOVA. Both repeated-measures ANOVA and mixed-effect analysis did not assume sphericity and were corrected by Geisser and Greenhouse epsilon hat method.

Figure experimental and analysis details

In Fig. 1a, schematic of a coronal section at 1.2 mm depicting viral injection of AAV1-FLEX-FLIM-AKAR into the NAc of *Drd1a-cre* and *Adora2a-cre* mice for D1R-SPN and D2R-SPN FLIM-AKAR measurements, respectively (left). An optical fibre was implanted 200 µm above the injection site (middle). A hybrid PMT and time-correlated single-photon counting were used to measure the fluorescence lifetime of FLIM-AKAR, which provides an estimate of the ratio between phosphorylated and non-phosphorylated FLIM-AKAR (right). Higher PKA activity results in a faster fluorescence decay (lower fluorescence lifetime) of the sensor (red).

In Fig. 1d, peak lifetime change was calculated from 100 s average around the peak.

In Fig. 2b, schematic of a sagittal section depicting viral expression and fibre implantation in *DAT-IRES-cre* mice. AAV9-hSyn-dLight1.1 and AAV1-hSyn-FLEX-NES-jRCaMP1b were injected into the NAc and VTA, respectively. Fibres were implanted 200 µm above the two injections sites. The NAc fibre was used to collect both dLight and DAN terminal jRCaMP signals. The VTA fibre was used to collect the DAN somatic jRCaMP signal.

In Fig. 2c–h, $\Delta f/f$ was normalized across all sessions for each mouse such that the 99th percentile response = 1.

In Fig. 2e, the mean of the normalized signal was calculated over 3 s starting at LED onset.

In Fig. 2f, the mean of the normalized signal was calculated over 3 s around the peak after the receptacle entry.

In Fig. 2g, the mean of the normalized signal was calculated over 0–3 s after time of expected pellet dispensing or 0–3 s after time of expected LED onset.

In Fig. 2i, R^2 was calculated from linear fit of each 20-s interval (−5 to 15 s with respect to the trigger zone entry) separately, averaged for each session, then averaged for individual mouse. R^2 between signal peaks was calculated from peaks above 1 s.d. of session data.

In Fig. 3a, schematic of a coronal section at 1.2 mm depicting injection of AAV9-hSyn-dLight1.1 and AAV1-FLEX-FLIM-AKAR into the NAc of *Drd1a-cre* or *Adora2a-cre* mice (left). Schematic of a coronal section depicting dual fibre photometry in which one fibre is used for intensity measurements of dLight fluorescence and the other for fluorescence lifetime measurements of FLIM-AKAR (right). The fibres were implanted 200 µm above injection sites.

In Fig. 3c top, mean of the normalized signal was calculated over 0–3 s after LED onset, 5.8–15.8 s after LED onset and 3 s around the peak after receptacle entry. 5.8s, the average delay between LED and pellet dispensing in regular rewarded trials.

In Fig. 3c middle, mean amplitude (ps) calculated over 0–40 s after LED.

In Fig. 3c bottom, mean amplitude (ps) calculated over 0–40 s after LED for success trials and 5.8–45.8 s after LED for failure trials. 5.8s, the average delay between LED and pellet dispensing in regular rewarded trials.

In Fig. 3e, mean of the normalized signal for dLight was calculated over 0–3 s after dispensing or time of expected reward. Average response of FLIM-AKAR (ps) was calculated from peak amplitudes after dispensing or time of expected reward.

In Fig. 3f, g, drug was intraperitoneally delivered 10 mins before recording began.

In Fig. 4a, schematic of a sagittal section depicting viral expression and fibre implantation. Left, AAV9-hSyn-dLight1.1 was injected into the NAc of *DAT-IRES-cre* mice. AAV1-hSyn-FLEX-ChrimsonR-tdTomato or AAV8-hSyn-SIO-stGtACR2-FusionRed was injected into the VTA for DAN activation or inactivation, respectively. Fibres were implanted 200 µm above the injection sites. Right, as on the left except AAV1-FLEX-FLIM-AKAR was injected to the NAc of *DAT-IRES-cre; Drd1a-cre* and *DAT-IRES-cre;Adora2a-cre* mice for, respectively, D1R-SPN and D2R-SPN FLIM-AKAR measurements.

In Fig. 4b–h, statistics for dLight and AKAR signal was performed on mean $\Delta f/f$ (0–3 s for activation, 0–10 s for inactivation) and mean lifetime (0–20 s for activation, 5–25 s for inactivation experiments), respectively. Average per mouse was calculated from 10 trials for optogenetic illumination and reward responses and 5 trials for reward-omission responses. Reward responses of SPN FLIM-AKAR were acquired in separate cohorts of untrained mice. Reward-omission responses of dLight and SPN FLIM-AKAR were acquired from separate cohorts that were trained on the behavioural task.

In Fig. 4b, stimulation parameters were 20 Hz train of 2-ms pulses at 14.3 mW. For food rewards, the response of dLight was aligned to its peak after receptacle entry and time shifted for maximum overlap with an optogenetic response.

In Fig. 4e, stimulation parameters were continuous illumination at 5 mW.

In Fig. 4h, intraperitoneal injection of drug was at least 10 min before recording. Stimulation parameters were 1 s illumination consisting of a 20-Hz train of 2-ms pulses at 14.3 mW (left) and 5 s continuous illumination at 2.5 mW (right).

Reporting summary

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

Data availability

All data (MATLAB data files) are available online via the public repository managed by Harvard Medical School (https://sharehost.hms.harvard.edu/neurobiology/?sabatini/DA_PKA). Any other relevant data are available from the corresponding author upon reasonable request.

Code availability

Custom MATLAB codes are available online via the public repository managed by Harvard Medical School (https://sharehost.hms.harvard.edu/neurobiology/?sabatini/DA_PKA).

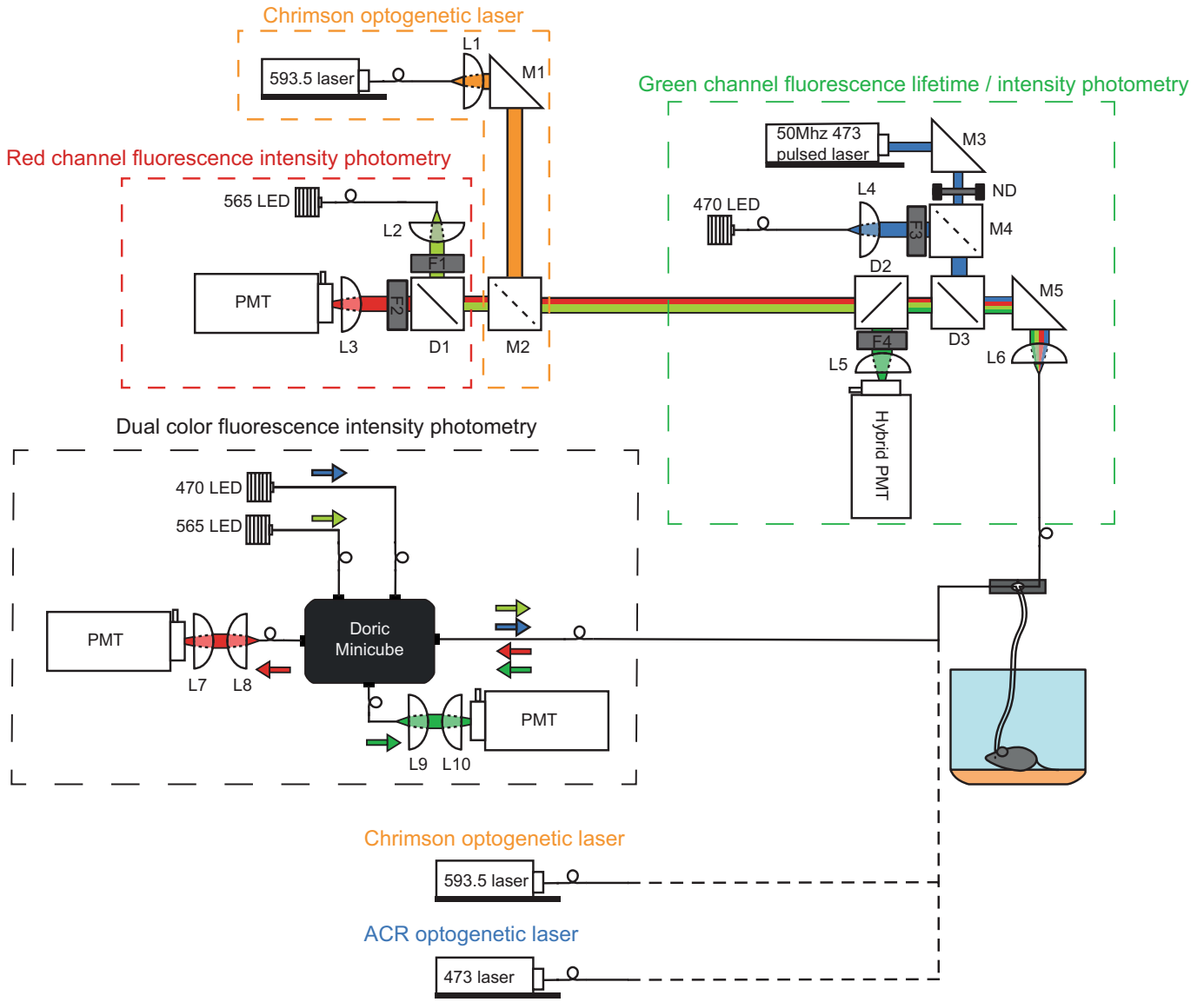
41. Gerfen, C. R., Paletzki, R. & Heintz, N. GENSAT BAC Cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron* **80**, 1368–1383 (2013).
42. Bäckman, C. M. et al. Characterization of a mouse strain expressing Cre recombinase from the 3′ untranslated region of the dopamine transporter locus. *Genesis* **44**, 383–390 (2006).
43. Lee, S. J., Escobedo-Lozoya, Y., Szatmari, E. M. & Yasuda, R. Activation of CaMKII in single dendritic spines during long-term potentiation. *Nature* **458**, 299–304 (2009).
44. Pnevmatikakis, E. A. et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* **89**, 285–299 (2016).
45. Motulsky, H. J. How to report the methods used for the mixed model analysis https://www.graphpad.com/guides/prism/8/statistics/stat_how-to-report-the-methods-used.htm (2020).

Acknowledgements This work was supported by NIH (B.L.S., U19NS113201 and R35NS105107; Y.C., F32DA035543; and L.T., U01NS013522 and U01NS090604), Howard Hughes Medical Institute (B.L.S.), Sackler Scholar Programme in Psychology (S.J.L.), Schuurman Schimmel van Outeren Stichting (B.L.) and Hendrik Muller fonds (B.L.). Graphical illustration was provided by Sigrid Knemeyer (sigrid@scistories.com).

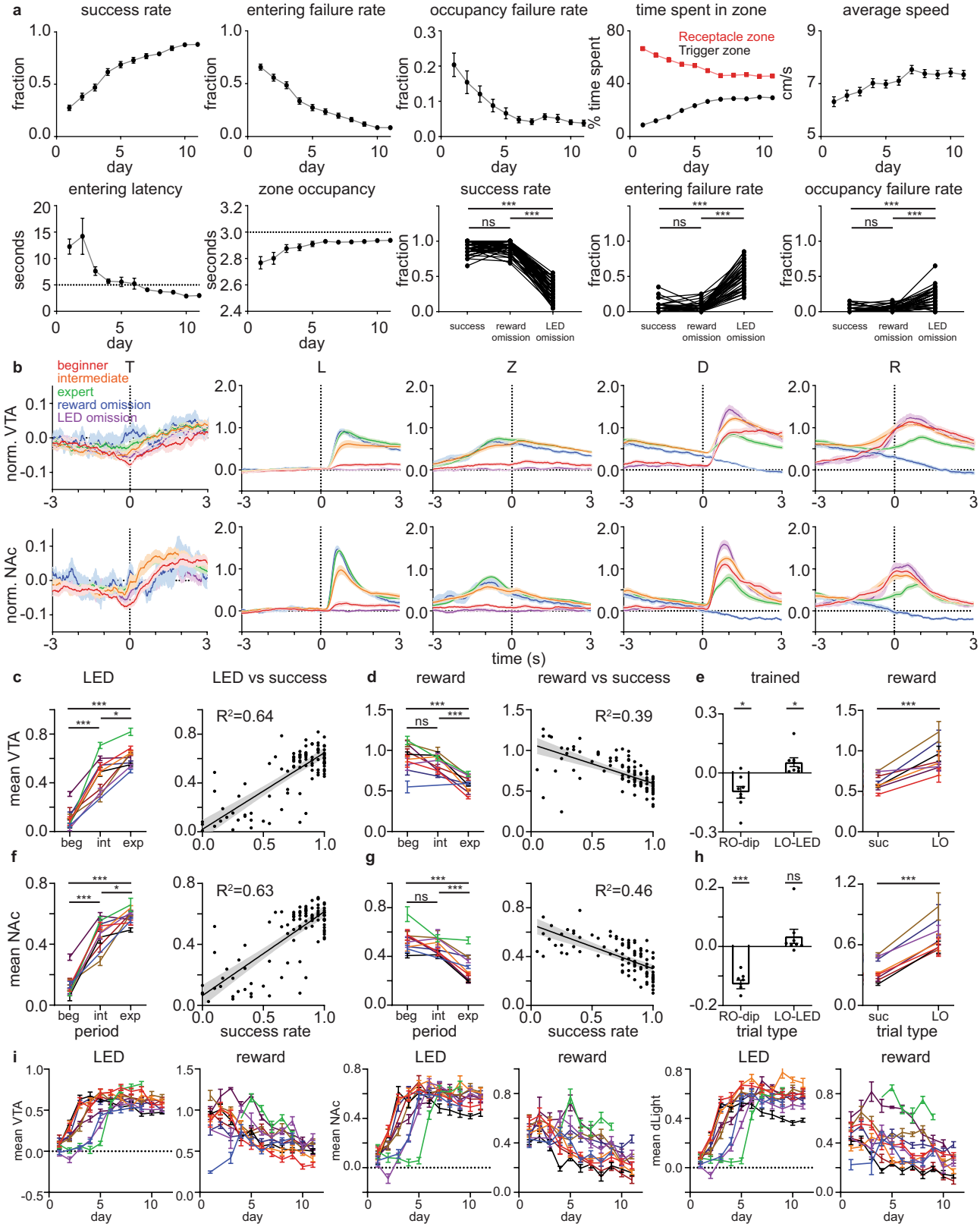
Author contributions S.J.L. designed the study, developed the photometry system, collected and analysed data and wrote the manuscript. B.L. collected and analysed data, and edited the manuscript. Y.C. conceived the study and edited the manuscript. T.P. and L.T. developed the dLight sensor and shared expertise. B.L.S. designed and supervised the study and wrote the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-03050-5.
Correspondence and requests for materials should be addressed to B.L.S.
Peer review information *Nature* thanks Michael Frank, Charles Gerfen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.

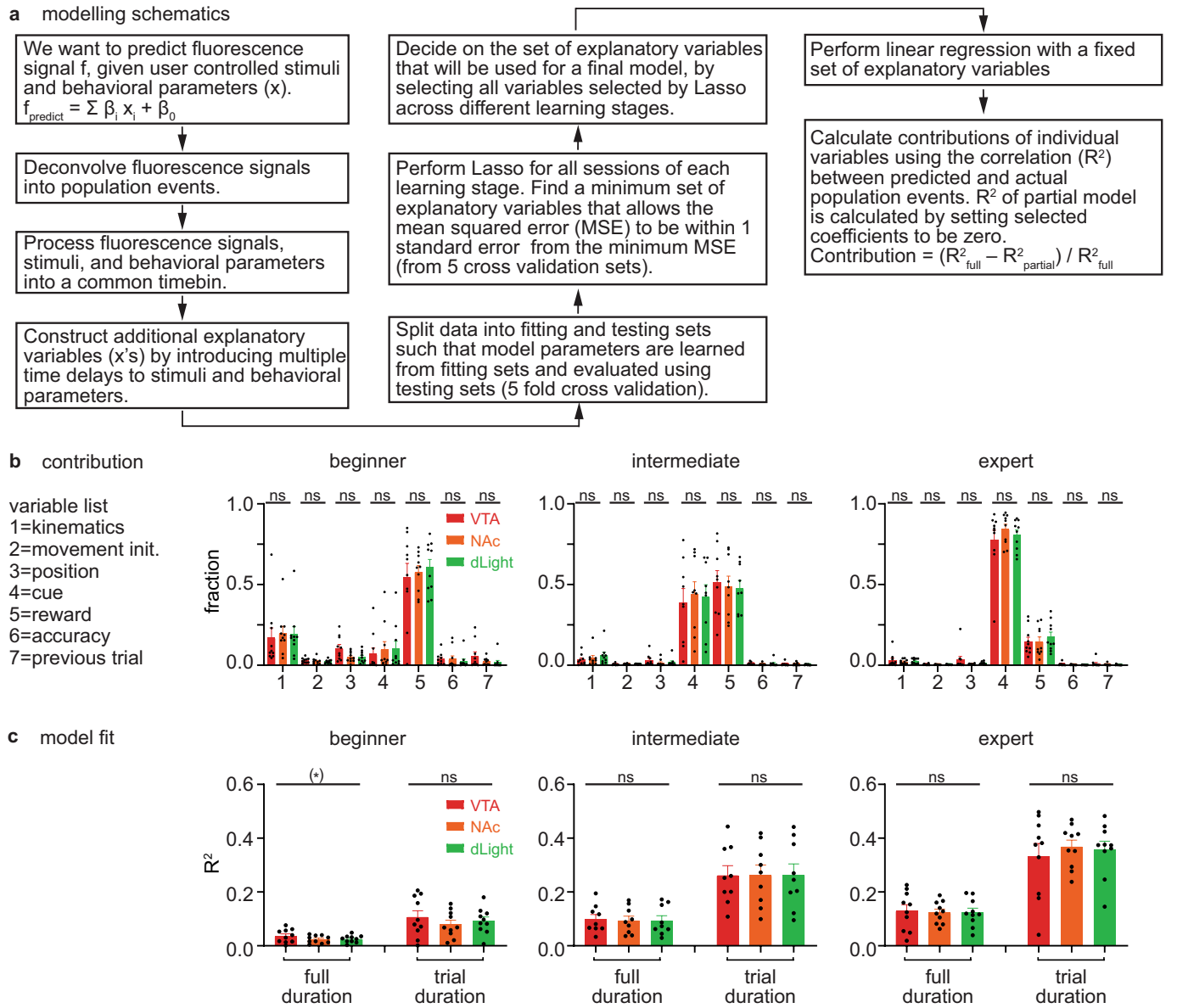


Extended Data Fig. 1 | Multipurpose photometry system for FLiP, fibre photometry and optogenetics. The system consists of two independent multicolour photometry units. The top photometry unit consists of three subcomponents used for: (1) red channel fluorescence photometry, (2) Chrimson optogenetic laser activation and (3) green channel fluorescence lifetime and intensity photometry. For (1), red channel photometry was accomplished using a fibre-coupled 565-nm LED (M565F3, Thorlabs) for excitation with output collimated in free-space by L2 and filtered by F1 (554/23, Semrock). Red fluorescence was separated from the excitation light by dichroic D1 (573LP, Semrock), filtered by F2 (630/60, Semrock) and focused onto a PMT (H10770(P)A-40, Hamamatsu) by L3. For (2), Chrimson optogenetic light was provided by a fibre-coupled 593.5-nm laser (SKU: YL-593-00100-CWM-SD-03-LED-F, Optoengine) with output collimated by L1 and combined with the red photometry path via M2, a mirror that can be inserted or removed, respectively, for Chrimson optogenetic stimulation or red channel photometry. For the green channel fluorescence lifetime measurement mode of (3), a 50-MHz 473-nm pulsed laser (BDS-473-SM-FBE, Becker and Hickl) was fed through a rotating neutral density filter for power adjustment, reflected by D3 (488LP dichroic, Semrock) and focused onto a patch cable by L6. Emission light was passed through D3, reflected by D2 (532LP dichroic, Semrock), filtered by F4 (517/22, Semrock) and focused by L5 to a high-speed hybrid PMT (HPM-100-07-Cooled, Becker and Hickl). The hybrid PMT was connected to a time-correlated single-photon counting board (SPC-830, Becker and Hickl) for fluorescence lifetime measurements. For the green channel fluorescence intensity measurement mode of (3), a fibre-coupled 470-nm LED (M470F3, Thorlabs) was collimated by L4, filtered by F3 (482/18, Semrock) and reflected by a removable mirror (M4); emission light was detected by a PMT (H7422-40, Hamamatsu). Alternatively, when fluorescence lifetime measurements were not needed, the bottom photometry unit was used. This simple 'dual-colour fluorescence intensity photometry' unit consists of 470-nm and 565-nm LEDs (Thorlabs), two PMTs (H10770(P)A-40, Hamamatsu) and a Doric Minicube (FMC5_E1(465-480)_F1(500-540)_E2(555-570)_F2(580-680)_S, Doric Lenses) that are connected by patch cables. For both photometry units, LEDs were driven by a digital signal processor system (RX8-5-12, Tucker-Davis Technology) for frequency modulation to carry out locked-in amplification of sensor signals detected by PMTs. In addition to the two main photometry units, a 593.5-nm laser (SKU: YL-593-00100-CWM-SD-03-LED-F, Optoengine) and a 473-nm laser (MBL-III-473, Optoengine) with independent patch cable connections were installed for Chrimson optogenetics and stGtACR2 optogenetics, respectively, for VTA DAN activity manipulation while monitoring the NAc.

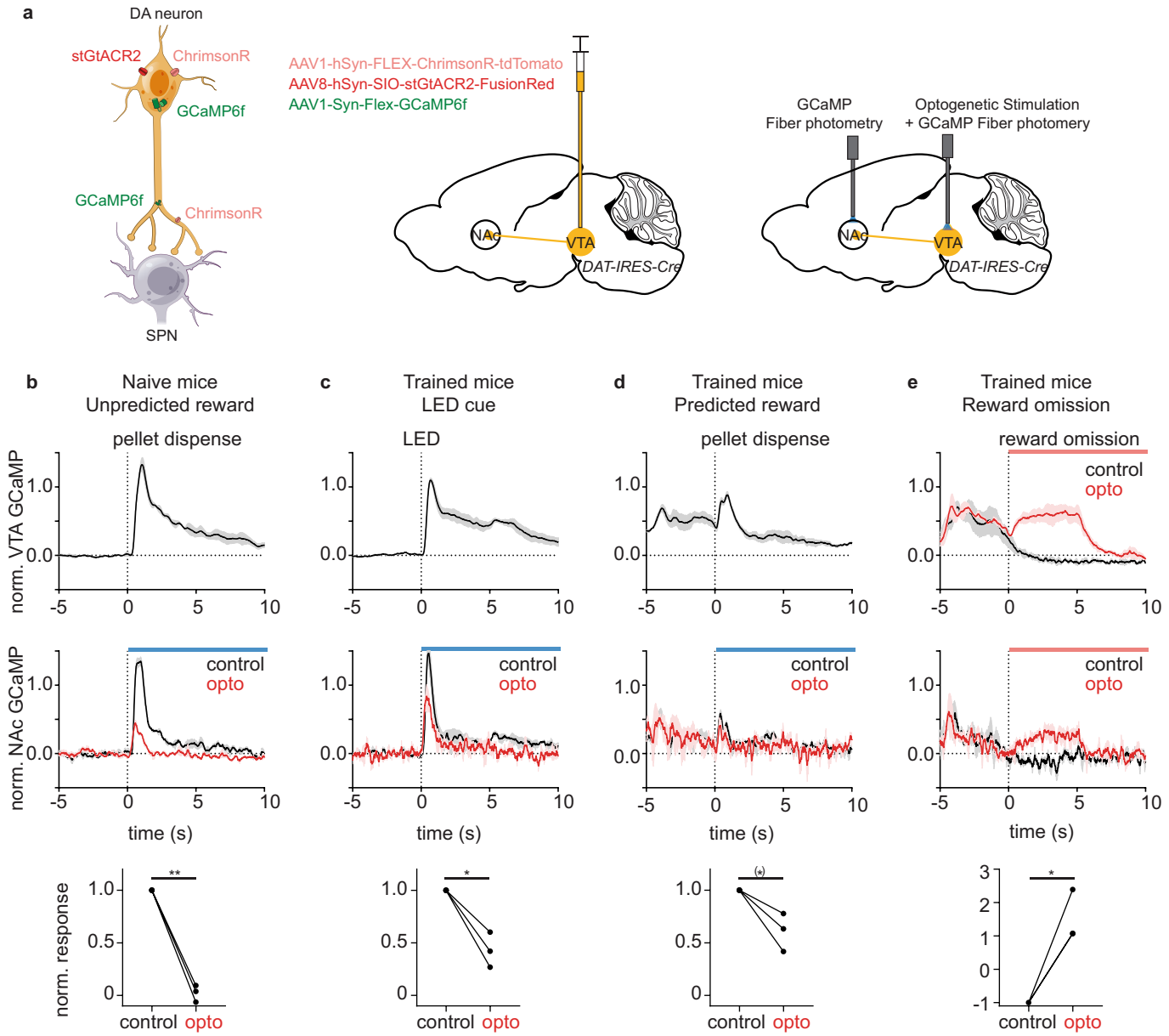


Extended Data Fig. 2 | See next page for caption.

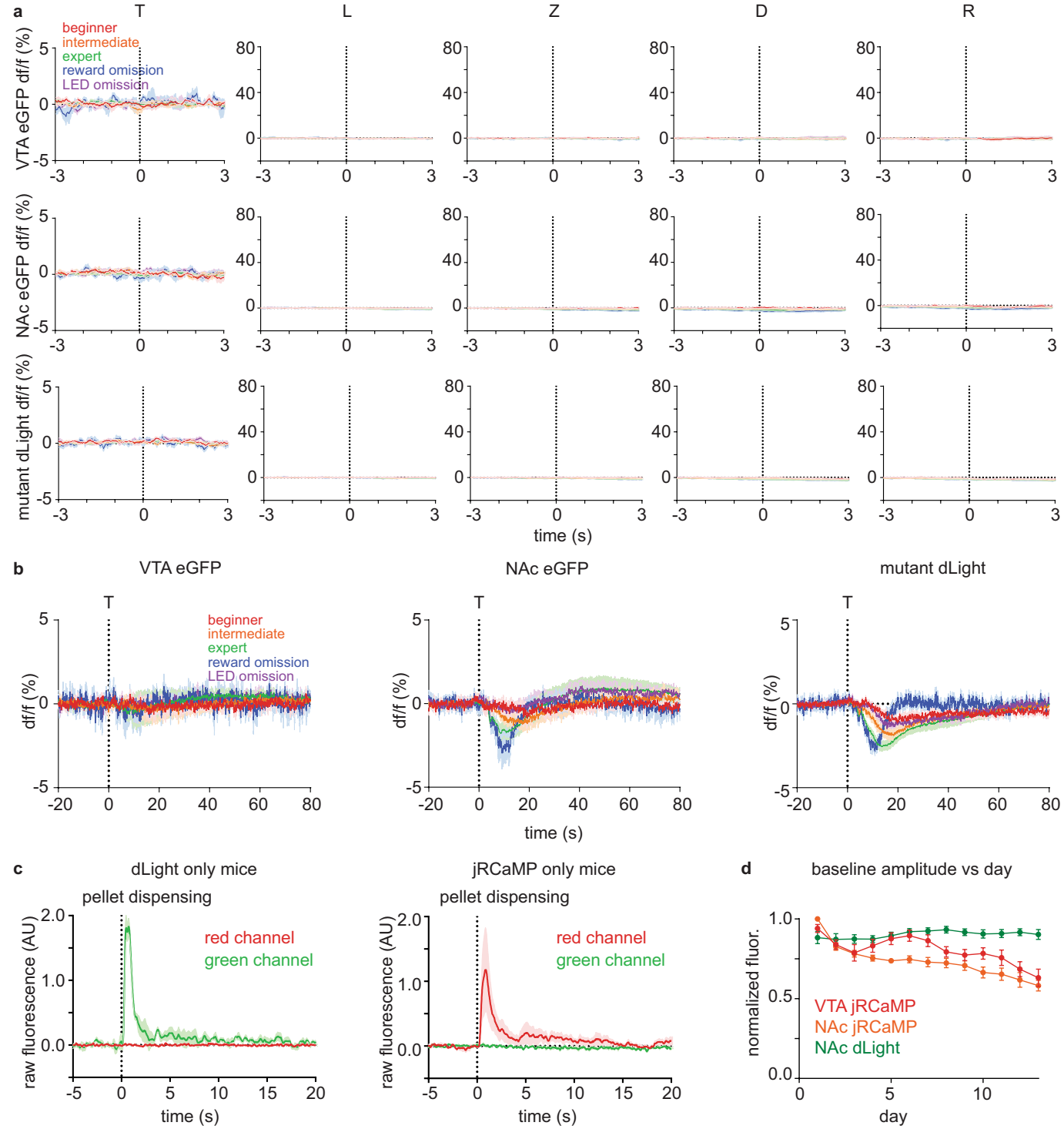
Extended Data Fig. 2 | Plasticity in DA release and DAN activity dynamics. **a**, Behavioural parameters demonstrating that mice are able to learn the visual-cue-guided operant conditioning described in Fig. 2a. Top: from the left, success rate (number of rewarded trials/total number of trials, one-way repeated-measures ANOVA, $F(6.298, 384.2) = 107.8$, $P < 0.0001$), entering failure rate (number of receptacle zone entering failure trials/total number of trials, one-way repeated-measures ANOVA, $F(5.782, 352.7) = 103.8$, $P < 0.0001$), occupancy failure rate (number of premature receptacle zone exit trials/number of receptacle zone entering success trials, one-way repeated-measures ANOVA, $F(4.253, 225.4) = 7.324$, $P < 0.0001$), time spent in zone (time spent in a zone/total session time, receptacle: one-way repeated-measures ANOVA, $F(3.171, 193.4) = 51.12$, $P < 0.0001$; trigger: $F(4.544, 277.2) = 110.9$, $P < 0.0001$), average speed (one-way repeated-measures ANOVA, $F(3.969, 242.1) = 15.26$, $P < 0.0001$). Bottom: from the left, entering latency (delay to enter the receptacle zone after the LED cue, one-way repeated-measures ANOVA, $F(1.760, 107.4) = 9.652$, $P = 0.0003$), zone occupancy (time spent in the receptacle zone after entering the zone during a trial, 3 s = maximum, one-way repeated-measures ANOVA, $F(5.229, 277.1) = 3.420$, $P = 0.0045$). Last three graphs depict success rate (one-way R repeated-measures ANOVA, $F(1.727, 72.54) = 668.6$, $P < 0.0001$), entering failure rate (one-way repeated-measures ANOVA, $F(1.328, 55.79) = 244.5$, $P < 0.0001$), and occupancy failure rate (one-way repeated-measures ANOVA, $F(1.295, 54.40) = 61.04$, $P < 0.0001$) comparisons for regular, reward-omission and rewarded LED-omission sessions of expert mice. $n = 64$ mice from all photometry behaviour experiments. Plotted as mean ± s.e.m. across mice and dots = mouse averages. $^{***}P < 0.001$ for Bonferroni-corrected post hoc comparisons. **b**, DAN activity across learning. The average responses for beginner, intermediate, expert, reward-omission (of expert mice) and rewarded LED-omission (of expert mice) trials are shown in red, orange, green, blue and purple, respectively. Dashed vertical lines indicate the behavioural time stamps (T = trigger zone entry, L = LED on, Z = receptacle zone entry, D = pellet dispensing, R = receptacle entry). Top: normalized $\Delta f/f$ of VTA jRCaMP signal showing VTA DAN soma activity. Bottom: normalized $\Delta f/f$ of NAc jRCaMP signal showing VTA DAN terminal activity. $n = 10$ mice. Plotted as mean ± s.e.m. across mice. **c**, VTA jRCaMP response (mean of normalized signal) to LED (mean of 0–3 s after LED cue) during training. Left, Individual mouse average plotted for different training periods (beginner, intermediate and expert) plotted as mean ± s.e.m. across trials (one-way repeated-measures ANOVA, $F(1.870, 14.96) = 73.76$, $P < 0.0001$). Right, daily average of LED response versus success rate, in which each dot represents a daily (session) measurement of a mouse plotted as linear regression fit ± its 95% confidence interval. **d**, VTA jRCaMP response to reward (mean of 3 s around the peak after receptacle entry) during training plotted as in **c** (one-way repeated-measures ANOVA, $F(1.835, 14.68) = 40.40$, $P < 0.0001$). **e**, VTA jRCaMP response in trained mice. Left, response to reward omission (RO-dip, mean of 3–6 s (shifted for slow soma jRCaMP signal) after expected time of pellet dispensing, one-sample t-test, $P = 0.013$) and to LED omission (LO-LED, mean of 0–3 s after expected time of LED onset, one-sample t-test, $P = 0.045$) plotted as mean ± s.e.m. across mice. Right, response to reward (mean of 0–3 s after pellet dispensing) in regular success (suc) and LED-omission (LO) trials (paired t-test, $P = 0.001$) plotted as in **c**. **f**, As in **c**, for NAc jRCaMP (one-way repeated-measures ANOVA, $F(1.986, 15.89) = 98.52$, $P < 0.0001$). **g**, As in **d**, for NAc jRCaMP (one-way repeated-measures ANOVA, $F(1.939, 15.51) = 33.95$, $P < 0.0001$). **h**, As in **e**, NAc jRCaMP response in trained mice. Left, response to reward omission (RO-dip, mean of 0–3 s after expected time of pellet dispensing, one-sample t-test, $P < 0.0001$) and to LED omission (LO-LED, mean of 0–3 s after expected time of LED onset, one-sample t-test, $P = 0.189$). Right, response to reward (mean of 0–3 s after pellet dispensing) in regular success (suc) and LED-omission (LO) trials (paired t-test, $P < 0.0001$). **i**, Daily average response of VTA jRCaMP, NAc jRCaMP and dLight response to LED and reward across training for individual mice. From the left, one-way mixed-effects analysis, $F(1.589, 13.35) = 25.07$, $P < 0.0001$; $F(2.339, 18.71) = 11.70$, $P = 0.0003$; $F(1.720, 14.45) = 25.75$, $P < 0.0001$; $F(2.984, 23.87) = 15.69$, $P < 0.0001$; $F(1.806, 15.17) = 35.41$, $P < 0.0001$; $F(2.835, 22.68) = 12.73$, $P < 0.0001$). Plotted as mean ± s.e.m.across trials for each mouse. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ for one-sample t-tests, and paired t-tests, Bonferroni-corrected post hoc comparisons. All t-tests are two-sided.



Extended Data Fig. 3 | Relationship between DA signal and behavioural parameters. a, Schematics of building a generalized linear model that relates user-controlled stimuli and behavioural parameters to fluorescence signals. In brief, there are three types of explanatory (independent) variables in the model. Continuous variables (speed, acceleration, rotation and position) continuously change their values as time passes. Event variables (movement initiation, cue, reward delivery and receptacle entry) are 0 except at a time point of an event when they temporarily change their value to 1. Whole trial variables (accuracy = 0 for current trial failure, 1 for current trial success; previous trial = 0 for previous trial failure, 1 for previous trial success) change their values in the beginning of a trial and stay constant until the next trial. **b**, Comparison of average variable contributions for VTA jRCaMP (red), NAc jRCaMP (orange) and dLight (green) for beginner (left), intermediate (middle), and expert (right) sessions. Contribution of each category was calculated by a method described in **a**. Kinematic variables include speed, acceleration and rotation variables. Other categories are assigned to an individual variable (a set of time-shifted variables). Mean contributions to three signals were compared by one-way repeated-measures ANOVA for each variable category. Plotted as mean ± s.e.m. across mice ($n = 10$ mice). **c**, Comparison of model fits for VTA jRCaMP (red), NAc jRCaMP (orange) and dLight (green) for beginner (left), intermediate (middle), and expert (right) sessions plotted as mean ± s.e.m. across mice ($n = 10$ mice). Model fit was estimated by the correlation between actual signals and predicted signals from the model. Left set of bars represent correlations during a full duration (about −40 to 80 with respect to the trigger zone entry). Right set of bars represent correlations during a trial duration (−5 to 15 s with respect to the trigger zone entry). Model fits for three signals were compared by one-way repeated-measures ANOVA. (*)$P < 0.10$ for one-way repeated-measures ANOVA (Bonferroni-corrected).

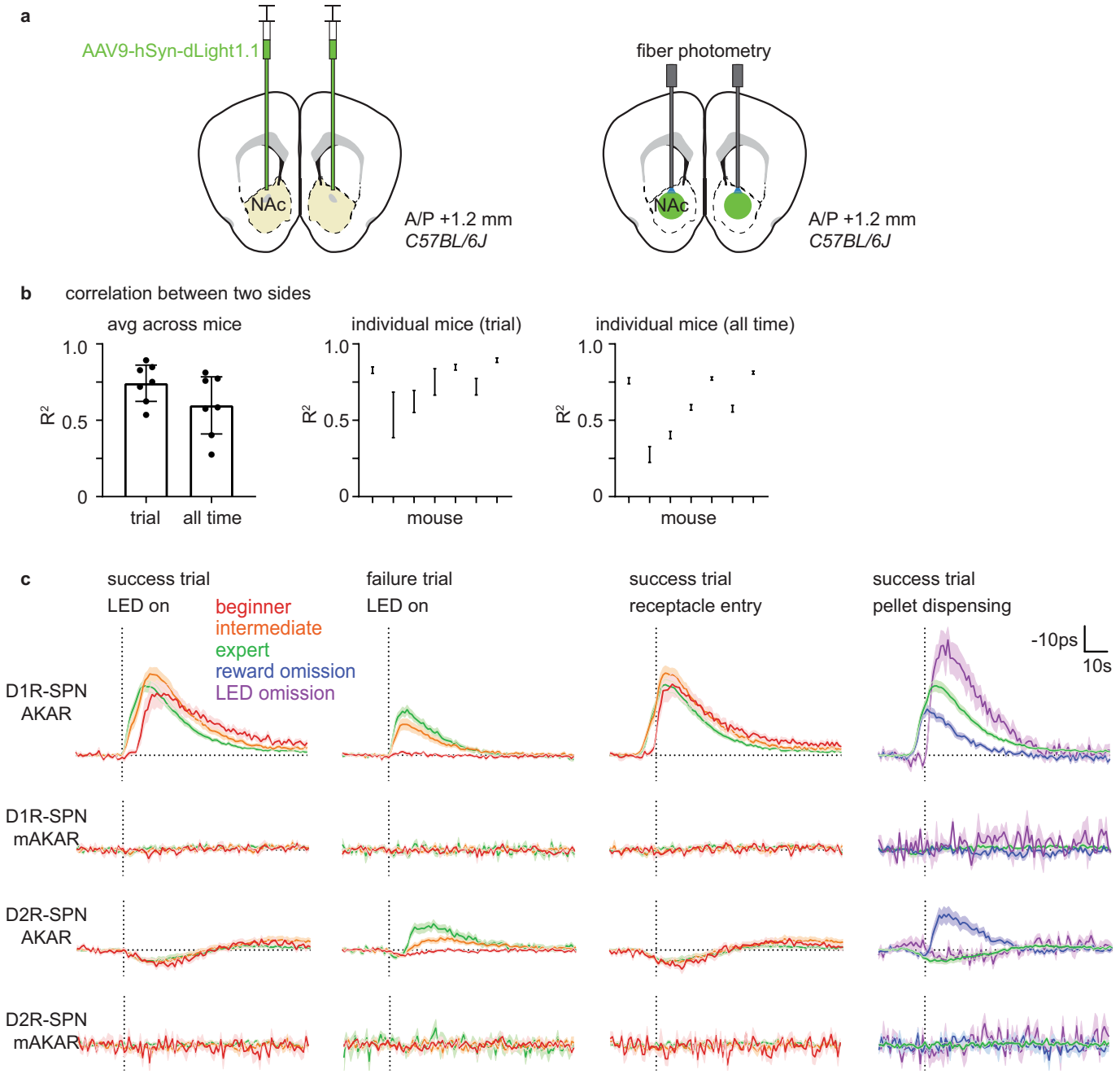


Extended Data Fig. 4 | Bidirectional modulation of DAN activity during behaviour. a, Schematic describing the experimental procedure. Left, expression of stGtACR2, ChrimsonR and GCaMP6f in DANs. Middle, injection of three viral vectors into VTA of a *DAT-IRES-cre* mouse. Right, the NAc fibre was used to collect terminal GCaMP signal. The VTA fibre was used to collect somatic GCaMP signal and to optogenetically activate (ChrimsonR) or inactivate (stGtACR2) DANs. **b**, VTA GCaMP (top) and NAc GCaMP (middle) response (normalized to 99th percentile of the single session) to unpredicted rewards in naive (untrained) mice. Mean NAc GCaMP signal normalized to the control response was compared between control and inactivation trials (bottom, paired *t*-test, $P = 0.002$). Dotted line represents time of pellet dispensing. **c**, VTA GCaMP (top) and NAc GCaMP (middle) response (normalized to 99th percentile of trained mouse sessions) to reward predictive LED cue in trained mice. Comparison of mean NAc GCaMP signal during control and inactivation trials (bottom, paired *t*-test, $P = 0.027$). Dotted line represents time of LED onset. **d**, VTA GCaMP (top) and NAc GCaMP (middle) response (normalized to 99th percentile of trained mouse sessions) to predicted reward (reward following LED cue) in trained mice. Comparison of mean NAc GCaMP signal during control and inactivation trials (bottom, paired *t*-test, $P = 0.065$). Dotted line represents time of pellet dispensing. **e**, VTA GCaMP (top) and NAc GCaMP (middle) response (normalized to 99th percentile of trained mouse sessions) to reward omission in trained mice. Comparison of mean NAc GCaMP signal during control and activation trials (bottom, 2 of 3 lines overlapping, paired *t*-test, $P = 0.029$). Dotted line represents time of expected reward delivery. (*)$P < 0.10$, *$P < 0.05$, **$P < 0.01$ for paired *t*-test. All graphs are plotted as mean ± s.e.m. across mice ($n = 3$ mice) and dot = mouse average. Average of signal (0–10 s) normalized to the control response was used for the comparison between control and optogenetic trials. The average response of each mouse was calculated from 5 or 6 trials. Blue and red bars indicate the periods of blue-laser illumination for stGtACR2 and red-laser illumination for ChrimsonR, respectively. VTA GCaMP signal could not be collected for blue-laser illumination period owing to optical crosstalk. All *t*-tests are two-sided.

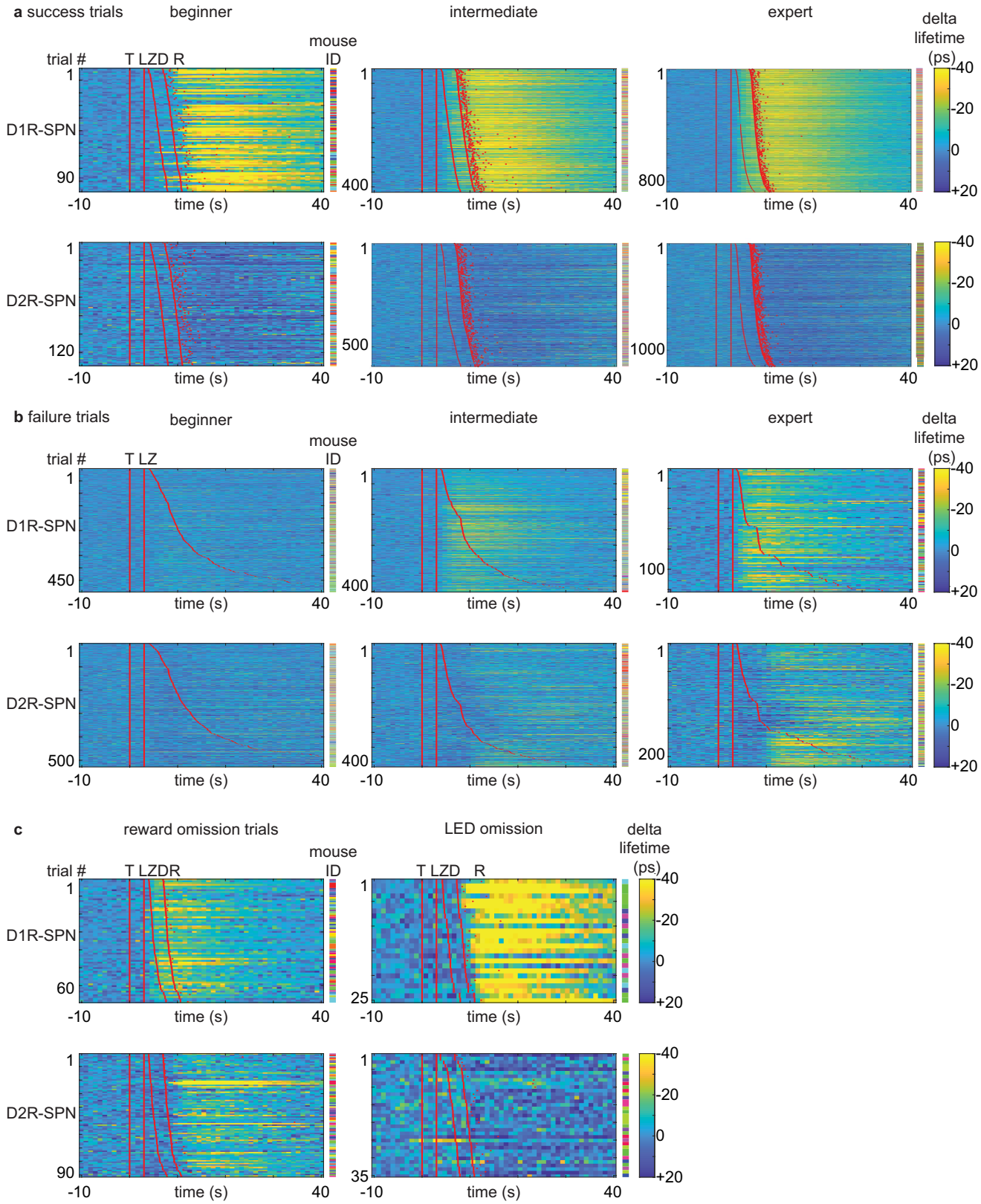


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Movement and optical artefacts cannot explain dLight and jRCaMP signal patterns. a, $\Delta f/f$ (%) of different controls. The average signals for beginner, intermediate, expert, reward-omission (of expert mice) and rewarded LED-omission (of expert mice) trials are shown in red, orange, green, blue and purple, respectively. Dashed vertical lines indicate the behavioural time stamps (T = trigger zone entry, L = LED on, Z = receptacle zone entry, D = pellet dispensing, R = receptacle entry). Top, $\Delta f/f$ (%) of eGFP signal from the VTA of *DAT-IRES-cre* mice ($n = 4$ mice) that were injected with AAV1-Cag-FLEX-eGFP into the VTA. Middle, $\Delta f/f$ (%) of eGFP signal from the NAc of *DAT-IRES-cre* mice ($n = 4$ mice) that were injected with AAV1-Cag-FLEX-eGFP into the VTA. Bottom, $\Delta f/f$ (%) of DA-binding mutant dLight (D103A mutation) signal from the NAc of C57BL/6J mice ($n = 8$ mice) that were injected with AAV9-hSyn-dLight^{D103A} into the NAc. **b**, $\Delta f/f$ (%) of different controls that are magnified in $\Delta f/f$ axis and demagnified in time axis. VTA eGFP (left, $n = 4$ mice), NAc eGFP (middle, $n = 4$ mice) and NAc mutant dLight (right, $n = 8$ mice) signal aligned to the time of trigger zone entry (dashed vertical line). There was a minor (compared to sensor responses) change in NAc eGFP and mutant dLight signal that develops across learning (possibly owing to haemodynamic effects). **c**, Test for the optical crosstalk between green and red spectrum for simultaneous dual-colour photometry for dLight and jRCaMP. Mice were given unexpected free food pellets, and signal was aligned to the time of pellet dispensing. Left, raw fluorescence signal in red and green spectrum from NAc of C57BL/6J mice ($n = 3$ mice, 10 trials prt mouse) injected with AAV9-hSyn-dLight1.1 into the NAc. Right, raw fluorescence signal in red and green spectrum from the NAc of *DAT-IRES-cre* mice ($n = 3$ mice, 10 trials per mouse) injected with AAV1-hSyn-FLEX-NES-jRCaMP1b into the VTA. **d**, Baseline (pretrial) raw fluorescence estimating the change in a signal strength due to photobleaching and viral expression change across days. Raw fluorescence was normalized by the maximum value of each mouse across all sessions. All graphs are plotted as mean ± s.e.m. across mice.

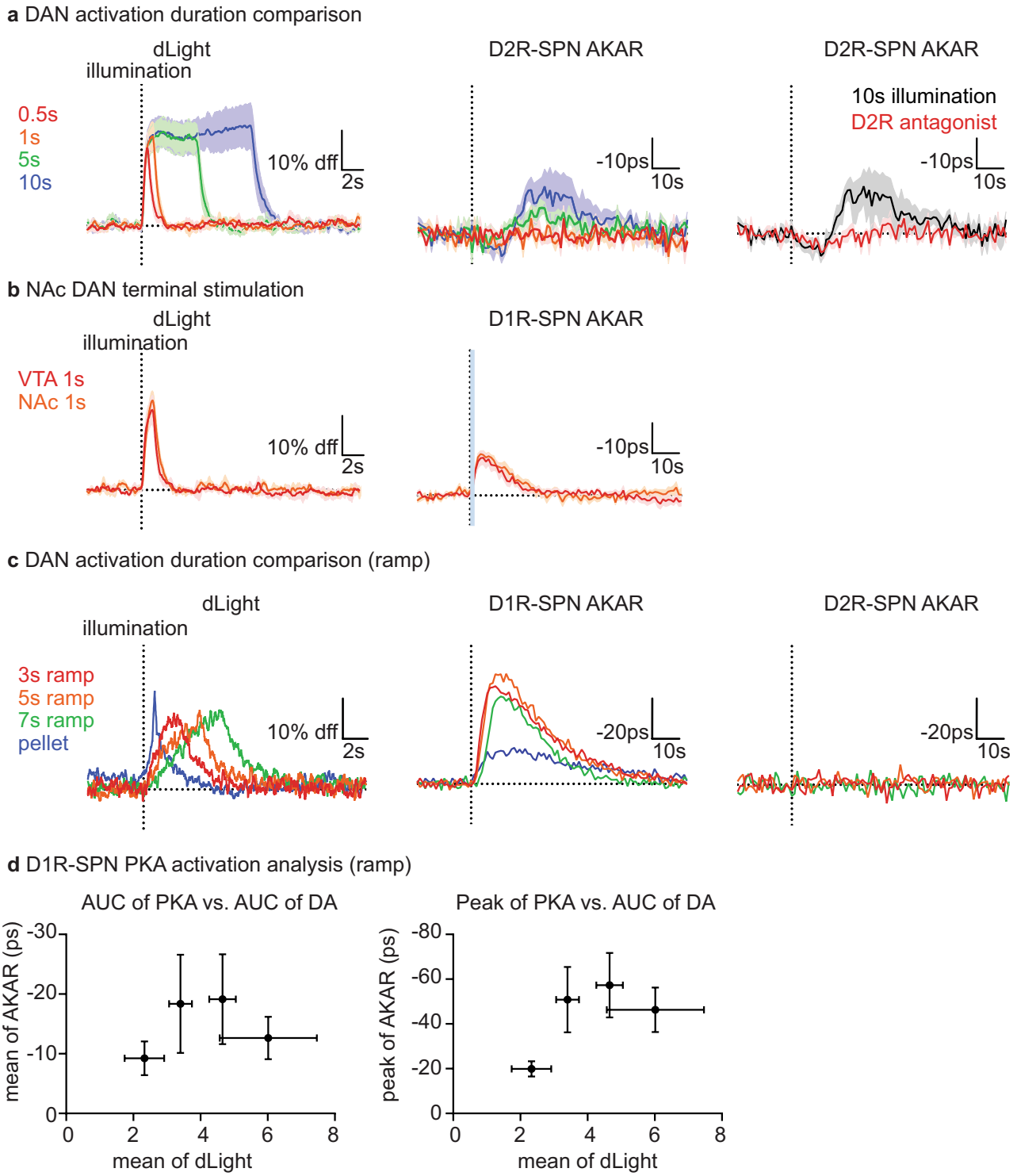


Extended Data Fig. 6 | Bilateral dLight measurement and mutant FLIM-AKAR control experiments. a, Schematic describing a strategy to measure DA level in both hemispheres. AAV9-hSyn-dLight1.1 was bilaterally injected into the NAc of C57BL/6J mice. Then, two optical fibres were implanted 200 µm above the injection sites in two hemispheres. **b**, Relationship between dLight signals from two hemispheres. Left, average correlation between two dLight signals (Δ*f*/*f*) during trial duration (−5 to 15 s with respect to the trigger zone entry) and all time (analysed in 20-s time window for each linear fit) plotted as mean ± 95% confidence interval across mouse averages (two-sided paired *t*-test, $P = 0.002$). Middle, correlation of individual mouse for trial duration plotted as mean ± 95% confidence interval across session averages for each mouse. Right, correlation of individual mouse for all time plotted as in middle. $n = 7$ mice. **c**, Comparison between FLIM-AKAR and FLIM-AKAR(T391A), which has a point mutation at the PKA phosphorylation site, signals. AAV1-FLEX-FLIM-AKAR or AAV1-FLEX-FLIM- FLIM-AKAR(T391A) was injected into the NAc of *Drd1a-cre* or *Adora2a-cre* mice for these experiments. From the top, D1R-SPN FLIM-AKAR (D1R-SPN AKAR), D1R-SPN FLIM-AKAR(T391A) (D1R-SPN mAKAR), D2R-SPN FLIM-AKAR (D2R-SPN AKAR), D2R-SPN FLIM-AKAR(T391A) (D2R-SPN mAKAR). From the left, signals were aligned to the time (dashed vertical line) of 'LED on' for success and failure trials separately, 'receptacle entry' for success trials and 'pellet dispensing' for success, reward-omission and LED-omission trials. Signals for beginner, intermediate, expert, reward-omission (of expert mice) and rewarded LED-omission (of expert mice) trials are shown in red, orange, green, blue and purple, respectively. Plotted as mean ± s.e.m. across mice. $n = 14$ mice (D1R-SPN AKAR), 7 mice (D1R-SPN mAKAR), 18 mice (D2R-SPN AKAR), 6 mice (D2R-SPN mAKAR).



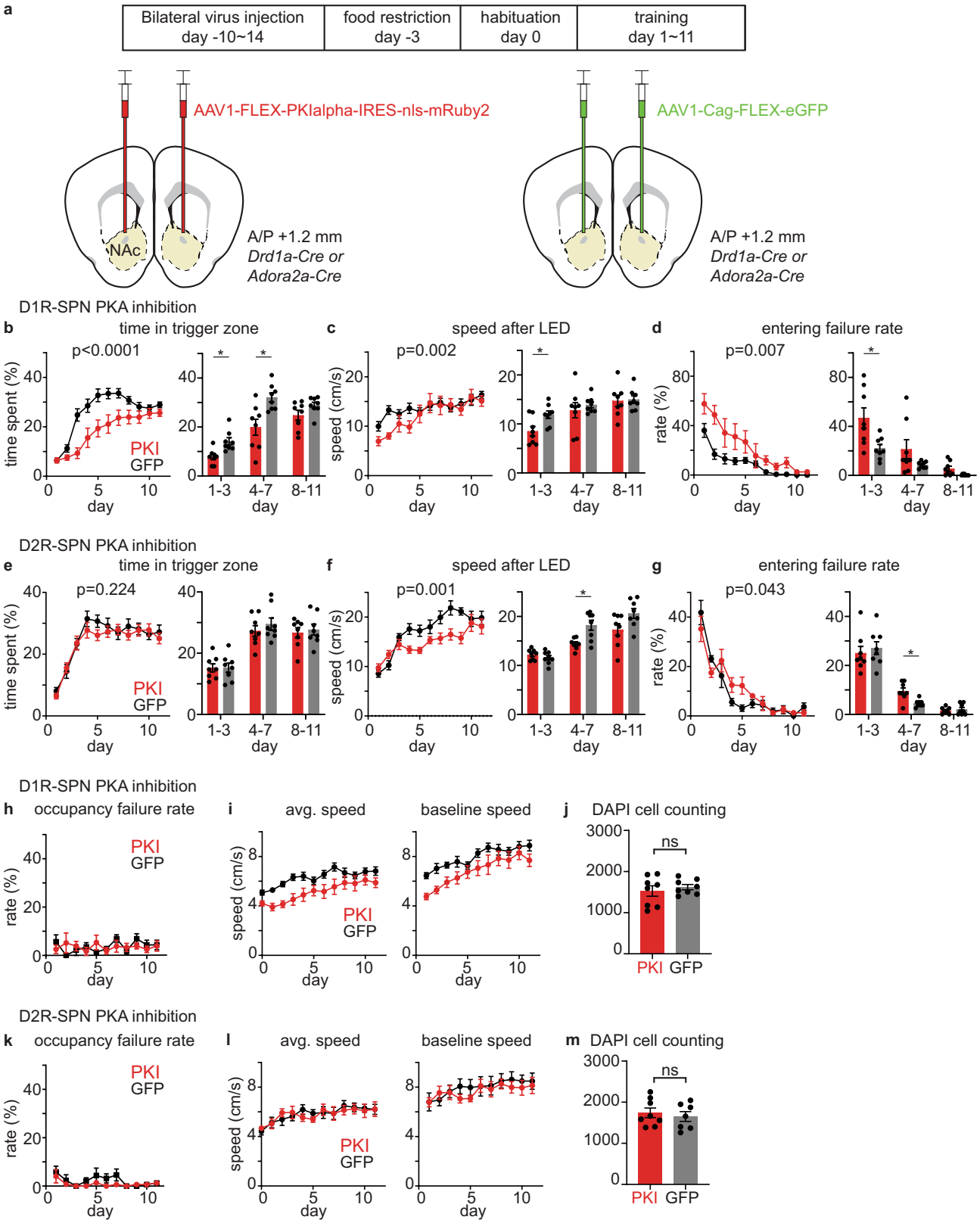
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Plasticity in patterns of PKA activity in SPNs during learning. a, Heat maps of SPN FLIM-AKAR response for success trials during learning. Each row represents an individual trial. Red lines or dots indicate behavioural time stamps (T = trigger zone entry, L = LED on, Z = receptacle zone entry, D = pellet dispensing, R = receptacle entry). Different colours in mouse ID columns represent different mice for an individual row. Top, D1R-SPN FLIM-AKAR responses of *Drd1a-cre* mice. *n* = 98 trials (beginner), 418 trials (intermediate), 873 trials (expert) from 14 mice. Bottom, D2R-SPN FLIM-AKAR responses of *Adora2a-cre* mice. *n* = 134 trials (beginner), 596 trials (intermediate), 1,152 trials (expert) from 18 mice. **b**, Heat maps of SPN FLIM-AKAR response for failure trials during learning. Plotted as in **a**. Top, D1R-SPN FLIM-AKAR responses of *Drd1a-cre* mice. *n* = 497 trials (beginner), 402 trials (intermediate), 122 trials (expert) from 14 mice. Bottom, D2R-SPN FLIM-AKAR responses of *Adora2a-cre* mice. *n* = 528 trials (beginner), 416 trials (intermediate), 218 trials (expert) from 18 mice. **c**, Heat maps of SPN FLIM-AKAR response for reward-omission trials (left) and rewarded LED-omission trials (right). Plotted as in **a**. Top, D1R-SPN FLIM-AKAR responses of *Drd1a-cre* mice. *n* = 69 trials (reward omission) from 14 mice, 25 trials (rewarded LED-omission trials) from 6 mice. Bottom, D2R-SPN FLIM-AKAR responses of *Adora2a-cre* mice. *n* = 91 trials (reward omission) from 18 mice, 35 trials (rewarded LED-omission trials) from 10 mice.



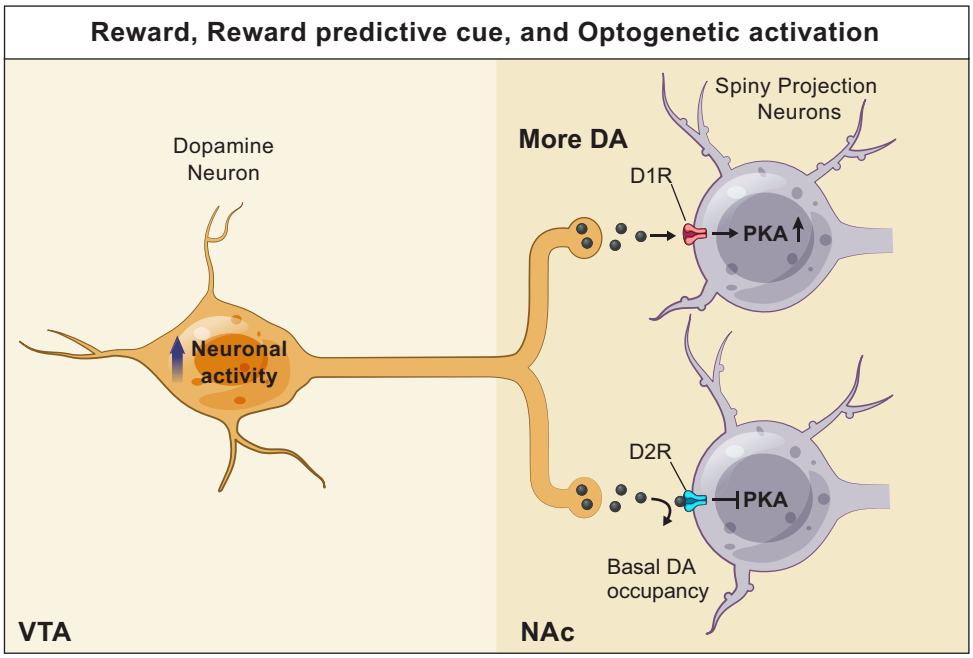
Extended Data Fig. 8 | See next page for caption.

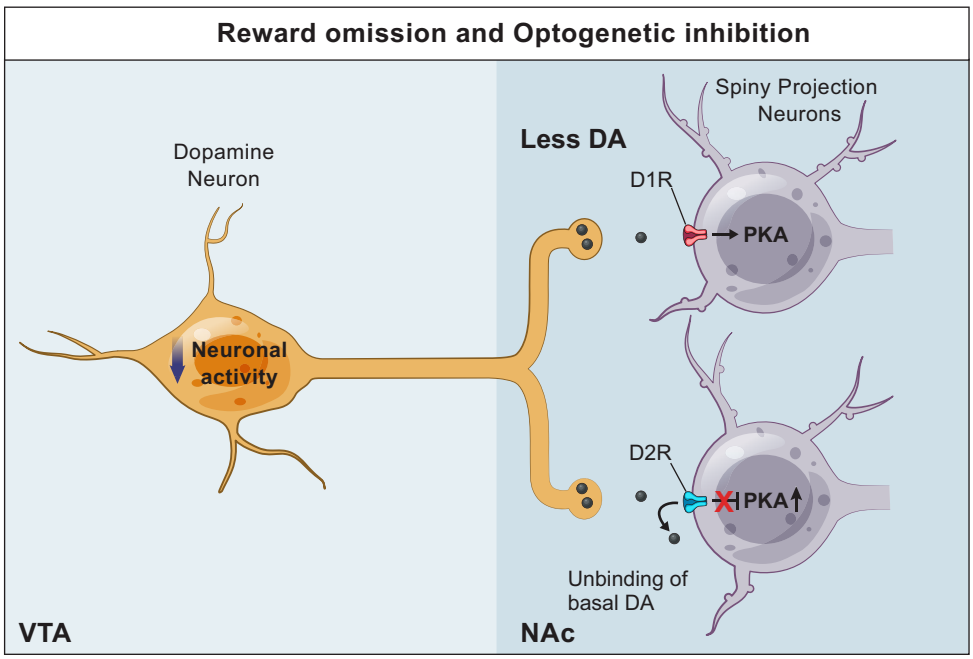
Extended Data Fig. 8 | Transient change in DAN activity is sufficient to modulate PKA activity in SPNs in the NAc. a, Left, dLight responses in *DAT-IRES-cre* mice ($n = 3$ mice) to DAN activation (20 Hz, 2-ms pulse width, 14.3 mW illumination) for different durations of illumination (red = 0.5 s, orange = 1 s, green = 5 s, blue = 10 s, one-way repeated-measures ANOVA, $F(1.064, 2.127) = 35.89$, $P = 0.023$). Middle, D2R-SPN FLIM-AKAR responses in *DAT-IRES-cre;Adora2a-cre* mice ($n = 4$ mice) plotted in the same way as the left (one-way repeated-measures ANOVA, $F(1.353, 4.059) = 4.140$, $P = 0.109$). Right, D2R-SPN FLIM-AKAR responses in *DAT-IRES-cre;Adora2a-cre* mice ($n = 4$ mice) to 10 s illumination without (black) and with (red) intraperitoneal injection of D2R antagonist at least 10 mins before recording (paired *t*-test, valley: $P = 0.338$, peak: $P = 0.033$). Statistics performed on mean dLight (0–10 s) and AKAR (end of illumination to ending time + 20 s except for valley estimation (0–10 s) for D2R antagonist experiment) signal. To test whether D2R-SPN PKA activity can respond to DAN activation at all, we activated DAN for 10 s, which increases DA levels far more than does a natural food reward response (**b**, left). This nonphysiological level of DA release results in a bidirectional modulation of D2R-SPN net PKA activity (**b**, right) with net PKA activity slightly decreasing and then increasing. However, given that D2R antagonist does not significantly affect the initial reduction in PKA activity, this reduction is unlikely to be D2R-mediated. On the other hand, the delayed activation of PKA was blunted by D2R antagonist, suggesting a contribution of indirect circuit mechanisms, such as modulation of the activity of D2R-expressing cholinergic interneurons in the NAc. **b**, dLight and D1R-SPN FLIM-AKAR responses to DAN terminal stimulation (20 Hz, 2-ms pulse width) in the NAc (red = VTA DAN stimulation for 1 s and 10.5 mW, orange = DAN terminal stimulation for 1 s and 7.7 mW). Left, dLight responses in *DAT-IRES-cre* mice ($n = 3$ mice, paired *t*-test, $P = 0.429$). Right, D1R-SPN FLIM-AKAR responses in *DAT-IRES-cre;Drd1a-cre* mice ($n = 6$ mice, paired *t*-test, $P = 0.597$). Statistics performed on mean dLight (0–3 s) and AKAR (end of illumination to ending time + 20 s) signal. **c**, Optogenetic induction of ramping DA level change and consequent PKA activity change in SPNs to different illumination duration. Left, dLight responses in *DAT-IRES-cre* mice ($n = 3$ mice) to ramping DAN activation for different ramping durations (red = 3 s, orange = 5 s, green = 7 s ramping activation) and to food reward (blue) (one-way repeated-measures ANOVA, $F(1.123, 2.246) = 4.260$, $P = 0.163$). To induce ramping DA level change, the frequency of stimulation was gradually increased from 24 Hz to 34 Hz for 3 s (at 10.5 mW), 16 Hz to 34 Hz for 5 s (at 6.1 mW), and 4 Hz to 30 Hz for 7 s (at 10.5 mW). Middle, D1R-SPN FLIM-AKAR responses in *DAT-IRES-cre;Drd1a-cre* mice ($n = 4$ mice) plotted in the same way as the left (one sample *t*-test on 7-s ramp, $P = 0.038$). Right, D2R-SPN FLIM-AKAR responses in *DAT-IRES-cre;Adora2a-cre* mice ($n = 4$ mice) plotted in the same way as the left (one sample *t*-test on 7-s ramp, $P = 0.779$). **d**, D1R-SPN PKA activation versus DA release analysis. Left, mean of D1R-SPN AKAR (0–80 s) versus mean of dLight (0–20 s) for different stimulations (reward, 3-s ramp, 5-s ramp and 7-s ramp). Right, peak of D1R-SPN AKAR (0–80 s) versus mean of dLight (0–20 s) for different stimulations. Each data point represents the average and the s.e. across mice ($n = 3$ mice for dLight, $n = 4$ mice for AKAR). All *t*-tests are two-sided. All graphs are plotted as mean ± s.e.m. (if shaded) across mice. Dashed vertical line = illumination onset. The average response of each mouse was calculated from 10 trials. Blue bars indicate the periods of laser illumination (NAc) for ChrimsonR during which accurate FLIM-AKAR measurements were not possible. Pellet response for dLight was aligned to the peak after receptacle entry and time-shifted so that the upward slope starts near 0 s. Pellet response for D1R-SPN AKAR was aligned to the receptacle entry. Statistics for **c** were performed on mean dLight signal (0–20 s) and mean AKAR signal (0–80 s).



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Selective PKA inhibition in SPNs slows learning. **a**, Schematic describing a strategy to investigate the effect of D1R-SPN or D2R-SPN PKA inhibition on behaviour. AAV1-FLEX-PKIalpha-IRES-nls-mRuby2 was injected into the NAc of *Drd1a-cre* or *Adora2a-cre* mice to selectively inhibit PKA in D1R-SPN or D2R-SPN, respectively. For control groups, AAV1-Cag-FLEX-eGFP was injected instead. Ten to fourteen days after surgery, mice were started on a behaviour schedule that includes 1 day of habituation (day 0) and 11 days of training (days 1 to 11). **b**, Effect of D1R-SPN PKA inhibition on the fraction of time spent in trigger zone (time spent in trigger zone/total session time). Daily average time (left, two-way repeated-measures ANOVA day × group interaction, $F(10, 140) = 5.565$, $P < 0.0001$) and multiday average (right, unpaired *t*-test, $P = 0.012, 0.017, 0.323$). **c**, Effect of D1R-SPN PKA inhibition on the speed after LED (average speed during 0–1.2 s after LED onset; 1.2 s is the minimum latency to enter the receptacle zone after LED cue). Daily average speed after LED onset (left, two-way repeated-measures ANOVA day × group interaction, $F(10, 140) = 2.923$, $P = 0.002$) and multiday average (right, unpaired *t*-test, $P = 0.046, >0.999, >0.999$). **d**, Effect of D1R-SPN PKA inhibition on the entering failure rate (number of entering failure trials/total number of trials). Daily average entering failure rate (left, two-way repeated-measures ANOVA day × group interaction, $F(10, 140) = 2.591$, $P = 0.007$) and multiday average (right, unpaired *t*-test, $P = 0.043, 0.367, 0.073$). **e**, Effect of D2R-SPN PKA inhibition on the time spent in trigger zone. Daily average time (left, two-way repeated-measures ANOVA day × group interaction, $F(10, 140) = 1.322$, $P = 0.224$) and multiday average (right, unpaired *t*-test, $P > 0.999, >0.999, >0.999$). **f**, Effect of D2R-SPN PKA inhibition on the speed after LED. Daily average speed after LED onset (left, two-way RM ANOVA day × group interaction, $F(10, 140) = 3.124$, $P = 0.001$) and multiday average (right, unpaired *t*-test, $P = 0.938, 0.011, 0.145$). **g**, Effect of D2R-SPN PKA inhibition on the entering failure rate. Daily average entering failure rate (left, two-way repeated-measures ANOVA day × group interaction, $F(10, 140) = 1.951$, $P = 0.043$) and multiday average (right, unpaired *t*-test, $P > 0.999, = 0.011, >0.999$). **h**, Occupancy failure rate (number of premature receptacle zone exit trials/number of receptacle zone entering success trials) of D1R-SPN PKA inhibition experiments (one-way repeated-measures ANOVA on GFP group, $F(3.213, 22.49) = 1.918$, $P = 0.153$; one-way repeated-measures ANOVA on PKI group, $F(2.424, 16.97) = 0.352$, $P = 0.747$). **i**, Left, average speed (total distance travelled/total session time, two-way repeated-measures ANOVA group effect, $F(1, 14) = 6.506$, $P = 0.023$). Right, average speed during baseline period (–20 s before trigger zone entry to trigger zone entry, two-way repeated-measures ANOVA group effect, $F(1,14) = 4.304$, $P = 0.057$). **j**, DAPI cell counting for D1R SPN PKA inhibition (unpaired *t*-test, $P = 0.517$). **k**, Occupancy failure rate of D2R-SPN PKA inhibition experiments (one-way repeated-measures ANOVA on GFP group, $F(2.293, 16.05) = 2.062$, $P = 0.155$; one-way repeated-measures ANOVA on PKI group, $F(1.598, 11.19) = 1.482$, $P = 0.264$). **l**, Left, average speed (two-way repeated-measures ANOVA group effect, $F(1, 14) = 0.011$, $P = 0.916$). Right, average speed during baseline period (two-way repeated-measures ANOVA group effect, $F(1, 14) = 0.211$, $P = 0.653$). **m**, DAPI cell counting for D2R SPN PKA inhibition (unpaired *t*-test, $P = 0.598$). $^*P < 0.05$ for Bonferroni-corrected unpaired *t*-tests. All *t*-tests are two-sided. Individual dots on the bar graph = a mouse. All graphs are plotted as mean ± s.e.m. across mice ($n = 8$ for each group).





Extended Data Fig. 10 | Model of DA action on PKA activity of SPNs. Overview of DA action on SPN PKA activity. Top, when DAN activity increases in response to a reward, reward predictive cue or optogenetic activation of DANs, more DA is released from DAN terminals in the NAc. This increase in DA level allows DA to bind to D1R, which increases the activity of adenylyl cyclase, the level of cAMP and—ultimately—the activity of PKA in D1R-SPNs. By contrast, the increase in DA level has a minimal effect on D2R, which is occupied by the basal level of DA. Bottom, when DAN activity decreases in response to a reward omission or optogenetic inhibition of DANs, DA release from DAN terminals in the NAc decreases below the baseline. This decrease in DA level has a minimal effect on D1R, which is not occupied by the basal DA. By contrast, the decrease in DA level allows the basal DA to unbind from D2R, which disinhibits PKA activity in D2R-SPNs.

nature research

Corresponding author(s): Bernardo Sabatini

Last updated by author(s): Oct 12, 2020

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.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☐	☒	A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Details are provided in methods section. Ethovision 11.5 and Bonsai 2.3 were used for behavioral control. Custom codes written in MATLAB 2012a (available online https://sharehost.hms.harvard.edu/neurobiology/?sabatini/DA_PKA) were used for photometry data collection.
Data analysis	MATLAB 2012a and 2019a were used for photometry data. OlyVIA 2.9 and Image J1.52i were used for histology data.

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. 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

Raw data are available online (https://sharehost.hms.harvard.edu/neurobiology/?sabatini/DA_PKA).

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No sample size precalculation was performed. Sample sizes are similar to related studies (Saunders, 2018; Mohebi, 2019) that manipulated or measured dopamine during behavior.
Data exclusions	Low-quality fiber photometry data of sessions (8 out of 128 sessions for dLight+jRCaMP experiments) in which baseline raw fluorescence was changed by more than 50% of the previous day's value was excluded.
Replication	We did not separately replicate the findings using new cohort of mice. However, we report all error bars and p-values for all our data, and our conclusions are based on statistically significant results.
Randomization	For PKI experiments, we randomly assigned mice to PKI and GFP group in a manner that each of a pair matched in age and sex was randomly assigned to a different group. We did not assign animals to different groups in other experiments.
Blinding	For PKI experiments, no blinding was attempted because the same experimenter carried out virus injection surgery and behavior experiments. We did not assign animals to different groups in other experiments. All behavior experiments were run by automated scripts.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

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

Methods

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

Animals and other organisms

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

Laboratory animals	Drd1a-Cre (B6.FVB(Cg)-Tg(Drd1-cre)EY262Gsat/Mmucd, 030989-UCD), Adora2a-Cre (B6.FVB(Cg)-Tg(Adora2a-cre)KG139Gsat/Mmucd, 036158-UCD), DAT-IRES-Cre (B6.SJL-Slc6a3tm1.1(cre)Bkmn/J, 006660), C57BL/6J (000664) mice, obtained from MMRRC-UC DAVIS, Jackson Laboratory, or bred in-house. For all experiments, age of mice was between 2-4 months, and both male and female mice were used in approximately equal proportion.
Wild animals	Our study did not involve wild animals.
Field-collected samples	Our study did not involve field-collected samples.
Ethics oversight	All animal procedures were approved by the Harvard Standing Committee on Animal Care following guidelines described in the US National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Note that full information on the approval of the study protocol must also be provided in the manuscript.