European Journal of Neuroscience, Vol. 49, pp. 726–736, 2019 doi:10.1111/ejn.13921

A silent eligibility trace enables dopamine-dependent synaptic plasticity for reinforcement learning in the mouse striatum

Tomomi Shindou, Mayumi Shindou, Sakurako Watanabe and Jeffery Wickens
Neurobiology Research Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1, Tancha, Onna-son, Okinawa 904-0412, Japan

Keywords: dopamine, learning, reinforcement, temporal difference

Abstract

Dopamine-dependent synaptic plasticity is a candidate mechanism for reinforcement learning. A silent eligibility trace – initiated by synaptic activity and transformed into synaptic strengthening by later action of dopamine – has been hypothesized to explain the retroactive effect of dopamine in reinforcing past behaviour. We tested this hypothesis by measuring time-dependent modulation of synaptic plasticity by dopamine in adult mouse striatum, using whole-cell recordings. Presynaptic activity followed by postsynaptic action potentials (pre–post) caused spike-timing-dependent long-term depression in D1-expressing neurons, but not in D2 neurons, and not if postsynaptic activity followed presynaptic activity. Subsequent experiments focused on D1 neurons. Applying a dopamine D1 receptor agonist during induction of pre–post plasticity caused long-term potentiation. This long-term potentiation was hidden by long-term depression occurring concurrently and was unmasked when long-term depression blocked an L-type calcium channel antagonist. Long-term potentiation was blocked by a Ca^{2+}-permeable AMPA receptor antagonist but not by an NMDA antagonist or an L-type calcium channel antagonist. Pre–post stimulation caused transient elevation of rectification – a marker for expression of Ca^{2+}-permeable AMPA receptors – for 2–4-s after stimulation. To test for an eligibility trace, dopamine was uncaged at specific time points before and after pre- and postsynaptic conjunction of activity. Dopamine caused potentiation selectively at synapses that were active 2-s before dopamine release, but not at earlier or later times. Our results provide direct evidence for a silent eligibility trace in the synapses of striatal neurons. This dopamine-timing-dependent plasticity may play a central role in reinforcement learning.

Introduction

Learning from positive reinforcement of previous actions underlies adaptive behaviour by increasing the frequency of actions that previously led to rewards. Dopamine – a neuromodulator released by positive reinforcement (Fiorillo *et al.*, 2003; Cheer *et al.*, 2007; Schultz *et al.*, 2008) – plays a key role in the underlying neural mechanisms. Dopamine-dependent synaptic plasticity in the corticostriatal pathway has been proposed as a cellular model for reinforcement learning (Reynolds *et al.*, 2001). In natural behaviour, however, reinforcement comes after the actions that led to reward, and at the cellular level, dopamine release is inevitably delayed relative to task-related neural activity. To bridge this delay between neural activity and dopamine release, a synaptic eligibility trace has been hypothesized (Miller, 1988; Izhikevich, 2007), in which phasically released dopamine strengthens synapses made eligible by previous activity. At the synaptic level, the eligibility trace may take the form of a silent synaptic trace activated by presynaptic and postsynaptic conjunction of activity, which later is converted into a long-term synaptic change by the action of dopamine. Despite its conceptual appeal, direct experimental evidence for a synaptic eligibility trace has been difficult to obtain. The aim of the research reported here was to test whether a silent eligibility trace exists in the corticostriatal synapses and to measure its temporal requirements.

Initial studies of synaptic plasticity in the striatum showed that a conjunction of presynaptic trains of high-frequency cortical stimulation with striatal postsynaptic depolarization caused long-term depression (LTD) (Calabresi *et al.*, 1992; Lovinger *et al.*, 1993). Application of dopamine reversed the depression that normally followed presynaptic and postsynaptic conjunction of activity, and caused long-term potentiation (LTP) supporting a three-factor synaptic modification rule (Wickens *et al.*, 1996). This potentiation was shown to be dependent on dopamine D-1 receptor activation *in vitro* (Kerr & Wickens, 2001) and *in vivo* (Reynolds *et al.*, 2001). The

Correspondence: J. Wickens, Neurobiology Research Unit, as above. Email: wickens@oist.jp

Received 1 November 2017, revised 2 March 2018, accepted 20 March 2018

Edited by Prof. Paul Bolam. Reviewed by Judith Walters, National Institute of Neurological Disorders and Stroke, USA; James Tepper, Rutgers University, USA; and Charles Wilson, University of Texas at San Antonio, USA.

All peer review communications can be found with the online version of the article.

reversal of the direction of synaptic change by dopamine receptor stimulation suggested that dopamine activated a molecular switch between LTD and LTP, and hinted at the possibility that LTD is itself a kind of eligibility trace, converted to LTP by application of dopamine. However, it is unknown whether LTD is a form of eligibility, or whether LTD and eligibility are independent states activated by similar conditions.

Prolonged trains of stimuli and bath-applied dopamine agonists, as used in earlier experiments, are unsuitable for detecting silent eligibility traces. The existence of a silent eligibility trace, and its temporal requirements in relation to presynaptic and postsynaptic activity, needs to be determined on a subsecond timescale. Stimulation protocols used to study spike-timing-dependent synaptic plasticity provide more precise control of timing of presynaptic and postsynaptic activity. Induction of synaptic plasticity in the striatum using spike-timing protocols has produced variable results (Thivierge *et al.*, 2007; Fino *et al.*, 2008; Pawlak & Kerr, 2008; Shen *et al.*, 2008; Shindou *et al.*, 2011). In these spike-timing-dependent plasticity inducing protocols, presynaptic spikes induced by electrical stimulation of afferent axons are paired with postsynaptic spikes caused by injection of suprathreshold current pulses. If presynaptic stimulation precedes postsynaptic spiking, the protocol is referred to as 'pre–post' pairing. Conversely, if presynaptic stimulation follows postsynaptic spiking, the protocol is referred to as 'post–pre'.

Experimentally, pre–post and post–pre stimulation protocols induce different forms of synaptic plasticity, depending also on the number of stimuli, spikes, the inter-stimulus intervals, and the rate and number of repetitions. Fino *et al.* (2008) reported spike-timing-dependent long-term depression (t-LTD) in corticostriatal synapses in response to pre–post spike timing, and spike-timing-dependent long-term potentiation (t-LTP) in response to post–pre timing in unidentified medium spiny neurons (MSNs). In contrast, Pawlak & Kerr (2008) reported t-LTP in response to pre–post spike timing and t-LTD in response to post–pre timing – the opposite of Fino *et al.* (2008) – in the presence of GABA blockers. Shindou *et al.* (2011) suggested there was a spine intracellular calcium concentration $[Ca^{2+}]_i$ threshold for induction of t-LTD in the corticostriatal pathway, mediated by the supralinear increase in $[Ca^{2+}]_i$ associated with pre–post induction protocols. In addition to differences in spike timing, functional differences among spiny projection neurons expressing different dopamine receptors (Gertler *et al.*, 2008) may modify spike-timing-dependent plasticity. In experiments that distinguished between dopamine D1 and D2 receptor expressing MSNs, Shen *et al.* (2008) found that intrastriatal stimulation using theta burst stimulation in pre–post sequence produced t-LTP in D1 neurons, which changed to LTD under dopamine-depleted conditions. However, Paille *et al.* (2013) found pre–post stimulation induced t-LTD in both D1-positive and D1-negative neurons. Yagishita *et al.* (2014) showed that co-occurrence of glutamate and dopamine stimulation was necessary for plasticity in corticostriatal synapses of D1-positive neurons, and potentiation did not occur if dopamine release was delayed until after glutamate activity ceased.

To obtain direct experimental evidence for a synaptic eligibility trace, it is necessary to establish the conditions for eligibility to occur. One possibility is that eligibility is established by a conjunction of presynaptic and postsynaptic activity, like that used in spike-timing-dependent plasticity studies. However, as noted above, a number of different findings have been reported concerning the effects of pre–post stimulation protocols in the corticostriatal pathway, including both t-LTD and t-LTP. Therefore, it is necessary to establish, as a baseline, reliable and reproducible effects of a standard induction protocol, to which phasic release of dopamine can be added. The conditions for induction of t-LTD provide a suitable starting point, based on the possibility that t-LTD might be converted to t-LTP by application of dopamine. Here, we report on an experimental test of the synaptic eligibility trace hypothesis based on measuring the effect of phasic dopamine release at different times relative to a conjunction of presynaptic and postsynaptic activity in the striatum. We first established that pre–post timing in the absence of dopamine caused t-LTD in D1 neurons and that this was associated with supralinear increases in dendritic spine $[Ca^{2+}]_i$ measured by two-photon microscopy. We then showed that entry of Ca^{2+} via L-type calcium channels was necessary for t-LTD, which could be converted to t-LTP by bath exposure to D1 agonists. We also found evidence for involvement of Ca^{2+}-permeable AMPA receptors in t-LTP. To precisely measure timing requirements, we used photolytic release of dopamine from caged dopamine (Lee *et al.*, 1996), in an explicit test of the silent synaptic eligibility trace hypothesis. We found that phasic dopamine release caused potentiation selectively at synapses that were active some seconds before dopamine release.

Materials and methods

Animals were handled in accordance with protocols approved by the Okinawa Institute of Science and Technology Animal Care and Use Committee ACUC# 2013-067-00.

Experiments were performed on male (age: 2–4 month) drd1a- and drd2-EGFP transgenic mice bred from Tg(Drd1-EGFP)118Gsat/Mmnc (MMRRC bioresource facility, University of Missouri/Harlan, USA) and Tg(Drd2-EGFP)118Gsat/Mmnc (MMRRC bioresource facility, University of North Carolina, USA) in which dopamine D1 or D2 receptor MSN subtypes expressed enhanced green fluorescent protein (EGFP) (Gong *et al.*, 2003).

Mice were deeply anaesthetized with sodium pentobarbital (80 mg/kg, WAKO) and then perfused transcardially for 2 min with cold modified-artificial cerebrospinal fluid (ACSF) containing the following (in mM): 50.0 NaCl, 2.5 KCl, 7.0 $MgCl_2$, 0.5 $CaCl_2$, 1.25 NaH_2PO_4, 25.0 $NaHCO_3$, 95.0 sucrose and 25.0 glucose and saturated with 95% O_2 – 5% CO_2. The brain was removed quickly and chilled. Slices (300 μm thick) containing the striatum were cut on a microtome (VT1000S, Leica) in an oblique plane, 45° rostral-up to the horizontal to preserve the corticostriatal projection. Slices were then incubated in oxygenated standard ACSF maintained at a temperature of 36 °C for 1 h. The standard ACSF had the following composition (mM): 120.0 NaCl, 2.5 KCl, 2.0 $CaCl_2$, 1.0 $MgCl_2$, 25.0 $NaHCO_3$, 1.25 $NaHPO_4$ and 15.0 Glucose. After incubation, a single slice was transferred to a recording chamber placed on the stage of an upright microscope and perfused (3–4–mL/min) with oxygenated ACSF at 30 °C. The remaining slices were kept in a holding chamber containing oxygenated ACSF at room temperature until required.

Electrophysiology

Whole-cell recordings from spiny projection neurons in dorsomedial striatum were made as previously described (Shindou *et al.*, 2011). Patch pipettes (2–4 MΩ) were filled with internal solution containing the following (mM): 115.0 K gluconate, 1.2 $MgCl_2$, 10.0 HEPES, 4.0 ATP, 0.3 GTP and 0.5% biocytin; pH 7.2–7.4. Glutamatergic excitatory postsynaptic potentials (EPSPs) were recorded in current clamp mode (Axon Instruments Multiclamp 700B). To test whether Ca^{2+}-permeable AMPA receptors (Shepherd, 2012) were transiently expressed during pre–post stimulation, we used an electrophysiological assay. To measure inward rectification, we used

voltage-clamp mode to measure excitatory postsynaptic currents (EPSCs) at different holding potentials interleaved with pre–post protocols (Plant *et al.*, 2006). We used the ratio of peak EPSC amplitude at negative (–80 mV) and positive (+40 mV) holding potentials as a measure of the rectification index (RI) at three time points (−2s, +2s and +4s). In these experiments, spermine was included in the internal solution.

Calcium imaging

Two-photon imaging of dendritic spines was performed with an upright microscope (FV1000MPE: Olympus) equipped with a 60 × objective (numerical aperture 0.9, Olympus), as previously described (Shindou *et al.*, 2011). Striatal spiny neurons were filled via the recording pipette with a combination of the Ca^{2+}-insensitive dye Alexa Fluor 594 (30 μM, Invitrogen) and Ca^{2+}-indicator Fluo-5F (300 μM, KD 2.3 μM, Invitrogen) for more than 20 min before imaging. Fluo-5F and Alexa Fluor 594 were excited using 850-nm light to monitor Ca^{2+} signals and spine morphology, respectively, and fluorescence emission was acquired at 400–570 nm (green) and 590–650 nm (red) for Fluo-5F and Alexa Fluor 594, respectively. Responsive spines on dendrites 30–60 μm from the soma were identified using frame scans. For quantitative measurements, line scans (500 Hz) were used. Intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) measures were based on the ratio of green fluorescence relative to baseline red fluorescence (R). We used the relation $\Delta R = R–R_0$ to estimate $[Ca^{2+}]_i$ changes and normalized to the maximal green fluorescence to red fluorescence (R_{max}), giving $\Delta R/R_{max}$. R0 was the averaged fluorescence over 50 ms immediately before stimulation. R_{max} was measured by delivering trains of action potentials (100–150 stimuli at 100 Hz) to saturate the indicator at the end of each patch clamp recording. Dual exponential fits to the fluorescence transients yielded the peak amplitude and the rising and decay phase time constants. Fluorescence traces are averages of three to five trials. Each spine measured was from a different cell, and all protocols were applied to each spine, permitting within-subject comparisons. Two-photon excitation at 720 nm (0.5 ms pulse) was used to cause photolysis of 4-methoxy-7-nitroindolinyl-caged-L-glutamate (MNI-glutamate, Tocris). MNI-glutamate (10 mM) was applied via a micropipette positioned above slice. Recorded neurons were identified during recording as D1 or D2 MSNs by their expression of EGFP colocalized with Alexa Fluor 594 from the recording electrode.

Synaptic plasticity

Plasticity experiments used spike-timing stimulation protocols described previously (Shindou *et al.*, 2011). A bipolar stimulating electrode was placed in the corpus callosum. Biphasic bipolar stimulation (0.01–0.5 mA, 200 μs duration) was used to evoke EPSPs with amplitudes of around 3 mV, range 1–5 mV. These electrical stimulation parameters were selected to avoid co-activation of dopaminergic fibres, and we confirmed using fast-scan cyclic voltammetry (FSCV) that with these parameters there was no-dopamine release. Baseline EPSPs were recorded for 10 min at 0.05 Hz stimulation rate. During pairing, single EPSPs were paired with a burst of three action potentials at 100 Hz in the postsynaptic cell. In pre–post protocols, the first of the three action potentials was evoked 10 ms after the EPSP onset. In post–pre protocols, the EPSP came 10 ms after the last of the three action potentials. Pairing was repeated 60 times at 10-s intervals. Neurons were identified as D1 or D2 MSNs by detection of EGFP in the recorded neuron during recording and subsequently confirmed to be MSNs by histological analysis.

Voltammetry

Dopamine levels were measured by FSCV (10 Hz triangular waveforms, −0.4 to 1.3 V versus Ag/AgCl electrode, 360 V/s) using carbon fibre microelectrodes (7 μm diameter with 150–200 μm exposed tip). Background-subtracted voltammetry signals were analysed offline using principal components regression (PCR) to isolate the dopamine component of the electrochemical currents (Wu *et al.*, 2001). Dopamine signals met established criteria and response to photolysis provided a well-characterized dopamine signal for reference.

Dopamine release in response to other spike-timing-dependent plasticity protocols was measured in two ways. To measure dopamine release in response to the stimulation protocol used by Shen *et al.* (2008), we used focal intrastriatal stimulation with a bipolar glass electrode (tip 10–15 μm diameter) pulled from glass capillary tubing with a septum dividing the lumen into two barrels (theta capillary glass) positioned 100–150 μm from the carbon fibre recording electrode. To measure release in the protocol used by Pawlak & Kerr (2008), we used single-pulse stimulation with a bipolar stimulating electrode positioned in the cerebral cortex and carbon fibre electrode in the striatum.

Phasic dopamine release

To measure the effects of phasic release of dopamine at different time points in relation to pairing of presynaptic and postsynaptic activity, we used ultraviolet (UV) flash photolysis of α-carboxy-*ortho*-nitrobenzyl (CNB) caged dopamine (Lee *et al.*, 1996). Dopamine was uncaged by a pulse (1 ms) of UV light from a Xenon light source (JML-C2 Xenon flash lamp system, Rapp OptoElectronics) directed onto the slice via a 40× objective lens. We measured uncaged dopamine using FSCV. This combination enabled us to produce dopamine pulses with precisely timed onset and time course mimicking the natural release of dopamine that occurs during reinforcement learning.

Experimental design and statistical analysis

Synaptic plasticity studies were based on between-groups comparisons of independent samples exposed to different pharmacological conditions. For statistical analysis of between-group differences, independent samples *t*-test was performed to detect significant differences between two groups, and two-way ANOVA was performed to detect interaction between treatment and time, with post hoc statistical tests where appropriate. Significance was taken at $P < 0.05$. Within-subjects comparisons in plasticity experiments, when used, were analysed using paired *t*-tests to compare EPSPs before and after pairing. Change in EPSP was measured using the averages of EPSPs recorded over the interval from 10 to 20 min after the end of the pairing period, relative to baseline. Group averages of responses were expressed as percentage change from the baseline EPSP amplitude. Data are presented as mean ± SEM. Statistical analyses were performed using Prism 4.0c (GraphPad) and SPSS21 (IBM).

Histology

After recording was completed, slices containing biocytin-loaded cells were fixed by immersion in 4% paraformaldehyde in 0.1 M phosphate buffer (PB) overnight at 4 °C. They were rinsed in PB for 30 min and then incubated in 15 and 30% sucrose for 30 min and 1 h, respectively. The slices were reacted with streptavidin-

conjugated Alexa Fluor 488 (1 : 1000) in PB containing 0.4% Triton X for 2 h at room temperature and mounted on slides. Labelled neurons were examined with a laser scanning confocal microscope (Zeiss LSM 780) using 20 × /0.8 Plan-Apochromat lens controlled by Zen 2012 software to confirm their identity as MSNs by the presence of spiny dendrites.

Results

Whole-cell records were made from 132 neurons in separate brain slices from 71 mice. Neurons were identified as D1 or D2 by expression of EGFP as observed during two-photon recording (Fig. 1a), and their identity as MSNs was confirmed by biocytin labelling and subsequent examination in fixed slices. Voltammetric records of extracellular dopamine release were obtained in 29 slices from seven mice.

In plasticity experiments, cells were exposed to either presynaptic-before-postsynaptic (pre–post) or postsynaptic-before-presynaptic (post–pre) pairing (Fig. 1b). Pre–post spike pairing caused spike-timing-dependent depression (t-LTD, Fig. 1c) of EPSPs in D1 MSNs ($-32.3 \pm 4.6\%$, $n = 10$, $t(9) = 6.49$, $P = 0.00011$, paired t-test) but not in D2 MSNs ($-12.8 \pm 7.4\%$, $n = 6$, $t(5) = 1.82$, $P = 0.13$, paired t-test). Conversely, post–pre protocols induced no change in either D1 ($-3.9 \pm 13.7\%$, $n = 8$, $t(7) = 0.32$, $P = 0.76$, paired t-test) or D2 ($-3.8 \pm 9.8\%$, $n = 6$, $t(5) = 0.61$, $P = 0.57$, paired t-test) MSNs.

Previous work has shown that a supralinear increase in spine $[Ca^{2+}]_i$ is associated with t-LTD of EPSPs in striatal MSNs (Shindou *et al.*, 2011). Supralinearity was identified when the $[Ca^{2+}]_i$ signals evoked by pre–post pairing exceed the algebraic sum of the $[Ca^{2+}]_i$ signals evoked separately by the EPSP caused by uncaging (uEPSP) and action potentials elicited as part of the spike-timing-dependent induction protocol. However, in Shindou *et al.* (2011), we could not identify D1 or D2 MSNs. In the current experiment, we definitively identified MSNs as D1 or D2 by their expression of EGFP. To understand t-LTD induction requirements related to cell subtype, we compared $[Ca^{2+}]_i$ transients evoked by glutamate uncaging causing uEPSPs in dendritic spines of D1 and D2 MSNs using two-photon stimulation and imaging (Fig. 2a). We focused on spines that were located as follows: D1 MSNs (total six cells), 49.97 ± 11.52 μm from soma, at second branch in five of six spines and at third branch in one spine; and D2 MSNs (total 6 cells), 48.14 ± 11.38 μm from soma, at second branch in all spines. Supralinear increases in $[Ca^{2+}]_i$ in dendritic spines (Fig. 2b) of D1 MSNs occurred uniquely with pre–post pairing (uEPSP = 0.13 ± 0.04; spikes = 0.18 ± 0.02; post–pre uEPSP + spikes, 0.25 ± 0.04; pre–post spikes + uEPSP, 0.39 ± 0.04, $n = 6$). Supralinear increases did not occur with either protocol in D2 MSNs (uEPSP, 0.09 ± 0.01; spikes 0.34 ± 0.06; post–pre uEPSP + spikes, 0.41 ± 0.07; pre–post spikes + uEPSP, 0.42 ± 0.05; $n = 6$). These data are summarized in Fig. 2c. Thus, a supralinear increase in spine $[Ca^{2+}]_i$ is associated with t-LTD of EPSPs in striatal D1 MSNs. However, neither supralinear increase in spine $[Ca^{2+}]_i$ nor t-LTD occurred with this stimulation protocol in D2 MSNs. As supralinear spine $[Ca^{2+}]_i$ increases and t-LTD were specific to D1 MSNs, the rest of this study focuses on D1 MSNs.

As shown in Fig. 3a, in D1 MSNs the L-type voltage-sensitive calcium (VSCC) antagonist nimodipine (10 μM) blocked t-LTD ($-4.2 \pm 7.1\%$, $n = 8$, $t(7) = 0.34$, $P = 0.74$, paired t-test). In contrast, the NMDAR antagonist APV (50 μM) did not block t-LTD ($-24.8 \pm 4.3\%$, $n = 8$, $t(7) = 5.81$, $P = 0.00066$, paired t-test).

Based on previous studies suggesting involvement of dopamine D1 receptors in plasticity (Kerr & Wickens, 2001; Reynolds *et al.*, 2001; Pawlak & Kerr, 2008; Shen *et al.*, 2008), we applied the D1

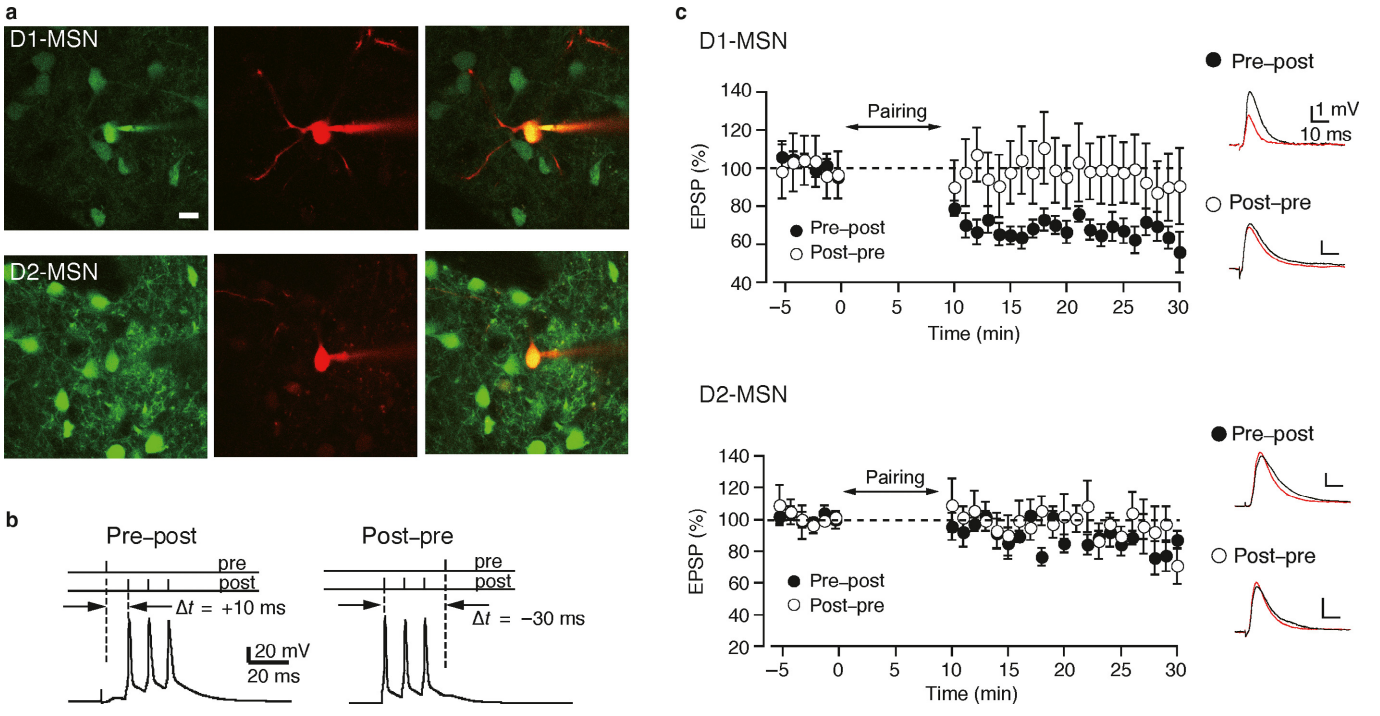

FIG. 1. Spike-timing-dependent plasticity in MSNs. (a) Identification of D1 (top row) and D2 (bottom row) MSNs. Left, EGFP-labelled neurons; centre, recorded cell labelled with Alexa Fluor 488; right, combined. (b) Pairing protocols. Left, pre–post pairing; right, post–pre pairing. Paired stimulation was repeated 60 times at 10-s intervals. (c) Pre–post stimulation caused t-LTD in D1 MSNs but not in D2 MSNs. Top, pre–post pairing compared with post–pre pairing in D1 MSNs. Bottom, same comparison in D2 MSNs.

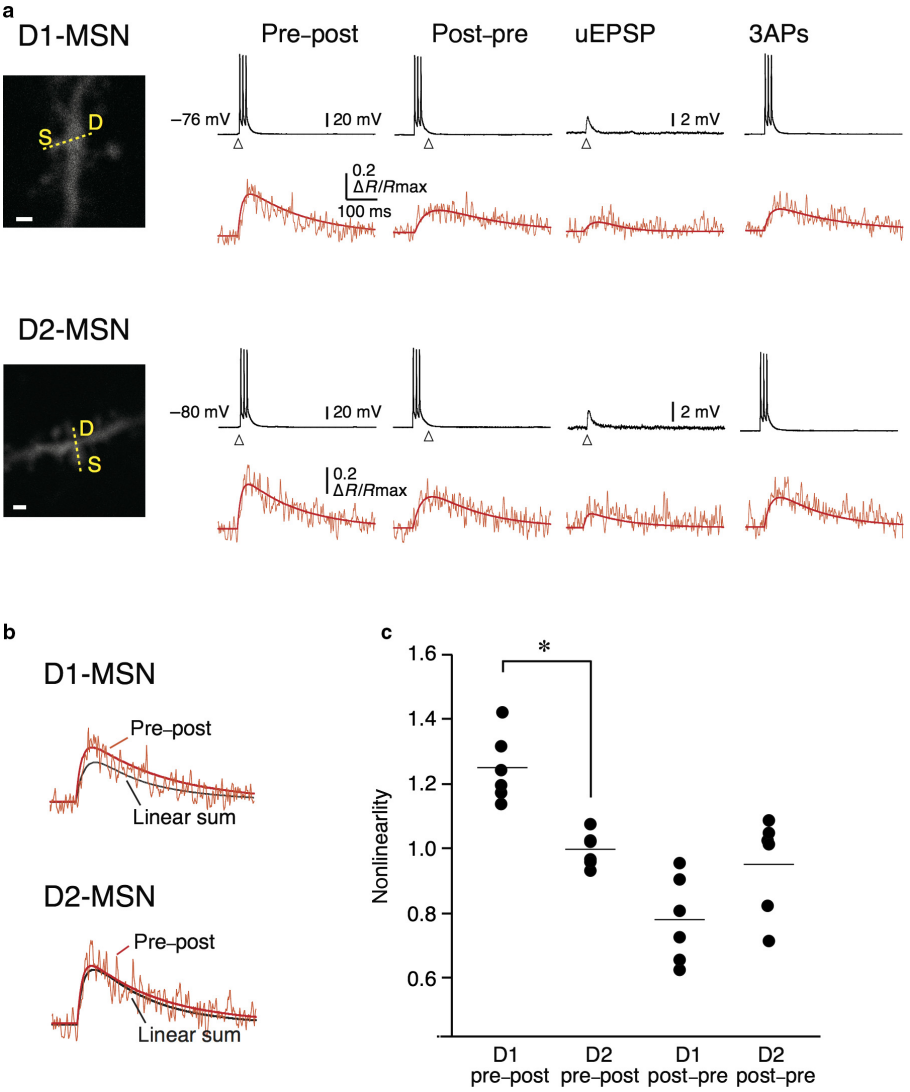

FIG. 2. Supralinear calcium influx associated with t-LTD. (a) Two-photon imaging of dendritic spines (S) and dendrites (D) in D1 and D2 MSNs showed $[Ca^{2+}]_i$ increase (red traces) in identified MSNs in response to pairing protocols (pre–post, post–pre), glutamate uncaging alone (uEPSP) and postsynaptic action potentials alone (3APs). (b) Supralinear dendritic spine $[Ca^{2+}]_i$ increase in D1-MSNs. Pre–post $[Ca^{2+}]_i$ increase exceeded the algebraic sum of pre and post alone, in D1 MSNs but not in D2 MSNs. (c) Plot shows significantly greater nonlinearity of pre–post $[Ca^{2+}]_i$ increase in D1 MSNs, relative to post–pre in D1 or D2 MSNs, of pre–post in D2 MSNs ($F_{1,20} = 20.17$, $P = 0.00022$, two-way ANOVA).

agonist SKF 81297 (1 μM) during stimulation that would normally induce t-LTD. The D1 agonist reversed t-LTD (−12.1 ± 3.1%, $n = 7$, $t(15) = 3.81$, $P = 0.002$, independent samples *t*-test, SKF vs. control, Fig. 3b, d) but did not switch LTD into LTP as previously reported using high-frequency stimulation paradigms (Wickens *et al.*, 1996). To test whether t-LTP occurred but was hidden by concomitant t-LTD, we blocked t-LTD with nimodipine (10 μM) during pre–post stimulation in the presence of SKF 81297 (0.5 μM). Under these conditions, t-LTP induction was observed in D1 MSNs (+53.2 ± 17.1%, $n = 7$, $t(6) = 3.31$, $P = 0.016$, paired *t*-test, Fig. 3b, d). Thus, dopamine D1 receptor activation throughout pre–post protocols caused t-LTP, which was unmasked when t-LTD was blocked.

To test for the involvement of NMDA receptors in t-LTP, we bath-applied the NMDA receptor antagonist APV (50 μM). We found that the NMDA receptor antagonist APV (50 μM) did not block t-LTP (+70.2 ± 56.2%, $n = 4$, $t(9) = 0.36$, $P = 0.73$, independent samples *t*-test, vs. nimodipine & SKF 81297 condition, Fig. 3c, d). This apparent contrast to previous studies in which NMDA receptor antagonists reduced t-LTP (Pawlak & Kerr, 2008; Shen *et al.*, 2008) was further investigated by measuring dopamine release in these protocols (reported below).

Labile memory traces that remain modifiable for a limited time interval have previously been associated with Ca^{2+}-permeable AMPA receptors (Shepherd, 2012). We hypothesized that synaptic eligibility traces might be associated with transient expression of these receptors. We found that t-LTP was blocked by the Ca^{2+}-permeable AMPA receptor antagonist, 1-naphthylacetyl spermine (NAS, 16.0 ± 12.0%, $n = 7$, $t(6) = 1.34$, $P = 0.23$, paired *t*-test, Fig. 3c, d), suggesting a transient eligibility trace mechanism linked to Ca^{2+}-permeable AMPA receptors.

Supralinear spine $[Ca^{2+}]_i$ increases (nonlinearity 1.25 ± 0.05 in control, $n = 6$) were blocked by the L-type voltage-sensitive calcium (VSCC) antagonist nimodipine (10 μM) (decreased to 1.03 ± 0.07 with nimodipine, $n = 7$, $t(11) = 2.57$, $P = 0.026$, independent samples *t*-test vs. control). While consistent with the block of t-LTD by

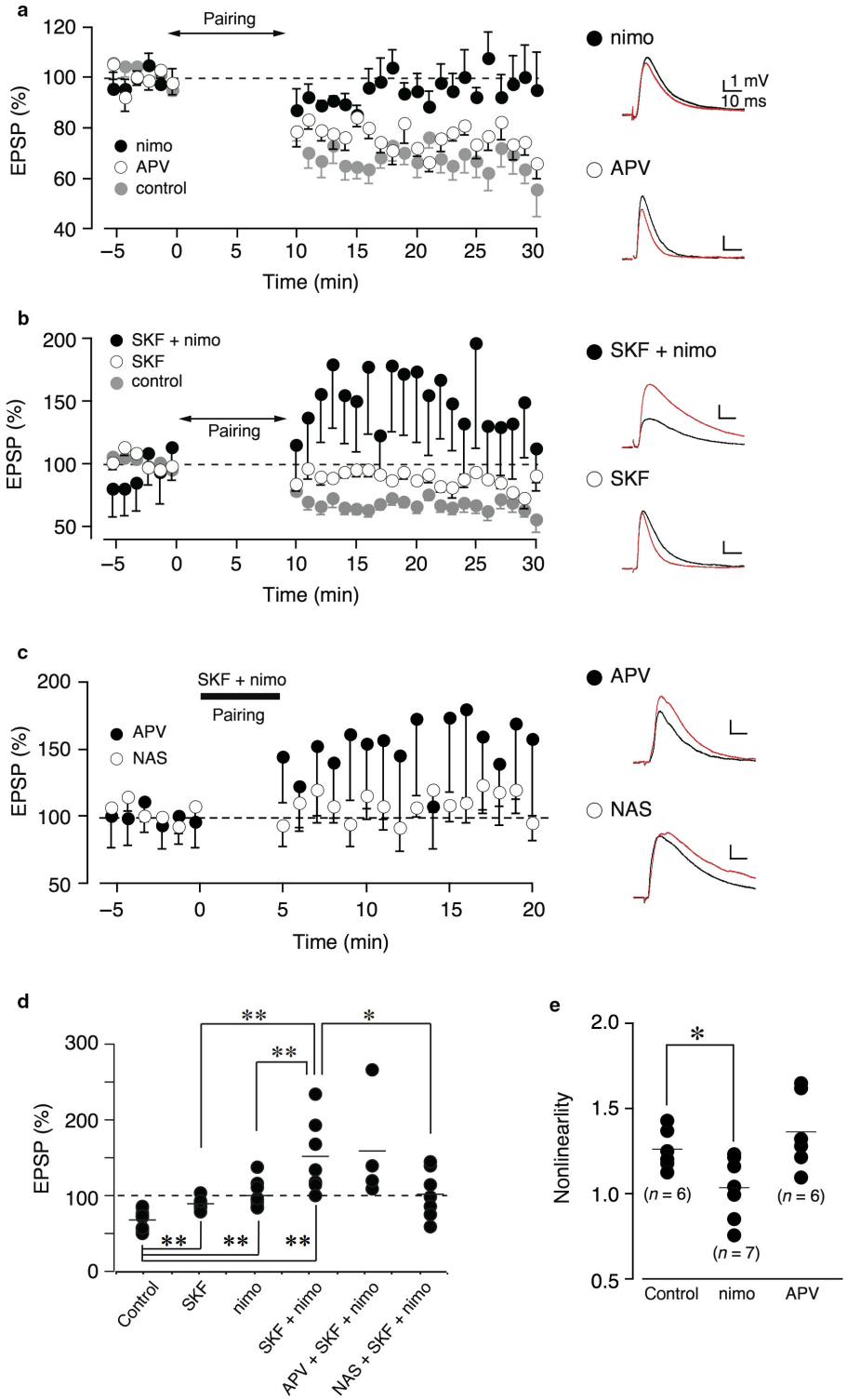

FIG. 3. Requirements for t-LTD and t-LTP. (a) Blocking L-type calcium channels with nimodipine (nimo) blocked t-LTD, but blocking NMDA channels with APV did not. (b) Pre–post stimulation plus dopamine D1 receptor agonist SKF 81297 (SKF) caused t-LTP when t-LTD was blocked by nimodipine. (c) Dopamine D1 receptor-dependent t-LTP was blocked by 1-naphthylacetyl spermine (NAS) but not by APV. (d) Summary data, all group averages: control vs. SKF: $t(15) = 3.81$, $P = 0.002$; control vs. nimo: $t(17) = 4.448$, $P = 0.00035$; control vs. SKF + nimo: $t(15) = 5.23$, $P = 0.0001$; SKF vs. SKF + nimo: $t(12) = 3.34$, $P = 0.006$; nimo vs. SKF + nimo: $t(14) = 2.96$, $P = 0.01$; SKF + nimo vs. NAS + SKF + nimo: $t(12) = 2.23$, $P = 0.046$; all independent samples *t*-tests. (e) Two-photon imaging of calcium shows that supralinear calcium influx in D1 MSNs (nonlinearity) was blocked by nimodipine. *, ** $P < 0.05$, $P < 0.01$.

nimodipine, our finding that t-LTP is not blocked by but rather unmasked by nimodipine shows that VSCC-mediated increases in spine $[Ca^{2+}]_i$ are not necessary for an eligibility trace. The NMDA receptor antagonist, APV (50 μM), did not block the supralinear $[Ca^{2+}]_i$ increase (1.36 ± 0.09, $n = 6$) seen with pre–post pairing in D1 MSNs ($t(10) = 1.008$, $P = 0.34$, independent samples *t*-test,

vs. control, Fig. 3e), consistent with lack of effect on t-LTD and t-LTP.

To test whether Ca^{2+}-permeable AMPA receptors (Shepherd, 2012) were transiently expressed during pre–post stimulation, we used an electrophysiological assay. Such receptors show inward rectification (Bowie & Mayer, 1995; Kamboj *et al.*, 1995; Koh *et al.*, 1995), allowing us to monitor rapid changes in the fraction of Ca^{2+}-permeable AMPA receptors by measuring inward rectification properties of EPSCs interleaved with pre–post protocols (Plant *et al.*, 2006). We found a significant increase in inward rectification index at 2s and 4s after pre–post stimulation ($F_{2,6} = 18.39$, $P = 0.003$, repeated measures ANOVA baseline vs +2s, 122.2 ± 31.5%, $n = 8$, $t(7) = 2.51$, $P = 0.04$, pairwise comparison with LSD; baseline vs +4s, 109.7 ± 6.8%, $n = 8$, $t(7) = 4.11$, $P = 0.005$, pairwise comparison with LSD, Fig. 4a).

The foregoing results indicate that the stimulation protocol used in the current experiments causes t-LTD in D-1 neurons, which is calcium-dependent, and requires calcium entry via L-type calcium channels, but not via NMDA channels. The same protocol – if combined with exposure to a D1 agonist throughout pairing activity – causes t-LTP, which requires calcium entry via Ca^{2+}-permeable AMPA receptor activation, but not NMDA channels. The t-LTP is

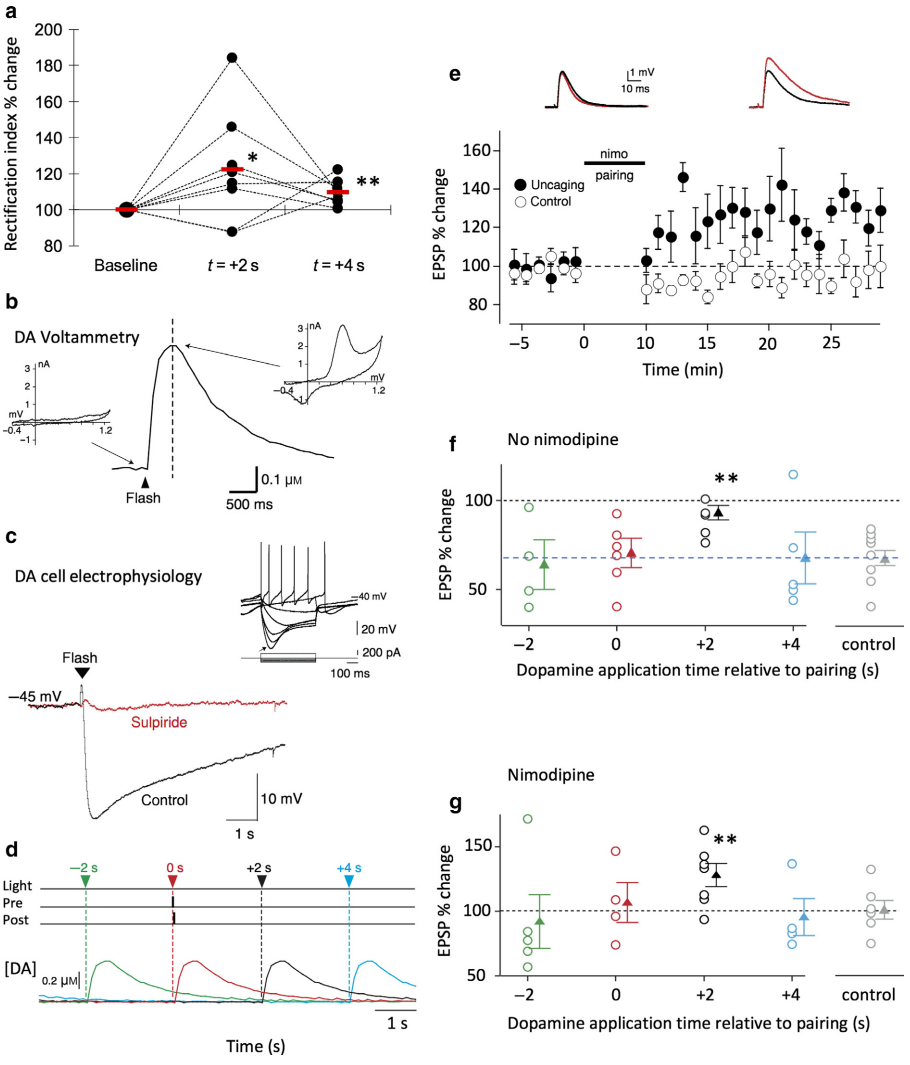

FIG. 4. Time-dependent modulation of spike-timing-dependent plasticity by dopamine supports eligibility trace hypothesis. (a) The rectification index (ratio of excitatory synaptic current at −80 mV to that at +40 mV under voltage clamp, normalized to baseline) indicates a transient increase in Ca-permeable AMPA receptors after pre–post pairing in D1 MSNs. Filled circles connected by dashed lines show individual neurons. Red bar shows mean. (b) Ultraviolet flash (arrowhead) causes phasic dopamine release by photolysis of caged dopamine (Lee *et al.*, 1996). Insets show background-subtracted fast-scan cyclic voltammograms at baseline (left) and peak response (right) with oxidation and reduction peaks for dopamine. Slope of trace is due to UV light effect on carbon fibre electrode. (c) Uncaged dopamine has functional effects. Dopamine release in response to UV light flash causes hyperpolarization of dopamine neuron (black trace, inset shows typical electrophysiological response to depolarizing current) that is blocked by a dopamine antagonist (red trace). (d) Timing diagram for phasic dopamine release evoked by UV uncaging (coloured triangles) in relation to presynaptic (pre) and postsynaptic (post) activity. Lower traces show predicted timecourse of phasic dopamine release evoked by flashes based on panel b. (e) Group average data showing that phasic dopamine release by uncaging 2s after pre–post stimulation causes t-LTP in D1 MSNs, in the presence of nimodipine. No change is seen in control condition (no uncaging). (f, g) In support of the eligibility trace hypothesis, t-LTP depends on timing of phasic dopamine release. (f) In the absence of nimodipine, group averages for dopamine release at −2, 0, +2 and +4s relative to pre–post conjunction of activity show t-LTP relative to no-dopamine, no-nimodipine control when uncaging occurs 2s after pre–post pairing, but no change relative to no-nimodipine control group at other time points. (g) In the presence of nimodipine, group averages for dopamine release at −2, 0, +2 and +4s relative to pre–post conjunction of activity show t-LTP relative to no-dopamine, nimodipine control group, when uncaging occurs 2s after pre–post pairing, but no change at other time points. *, ** $P < 0.05$, $P < 0.01$.

not blocked by but rather unmasked by blockade of L-type calcium channels. This background enabled us to test whether phasic release of dopamine at different times relative to pre–post conjunction of activity would produce t-LTP similar to that produced by exposure to a D1 agonist throughout pairing activity.

Photolytic uncaging of CNB caged dopamine (Lee *et al.*, 1996) by UV flash photolysis produced dopamine pulses with precisely timed onset and timecourse mimicking the natural release of dopamine that occurs during reinforcement learning. Using FSCV, we confirmed reliable and repeatable production of the temporal concentration profile of dopamine within the slice (peak concentration approximately 0.5 μM occurred 400 ms after the UV flash, Fig. 4b, d). We also confirmed physiological effects of uncaging dopamine on dopamine neurons in brain slices of the substantia nigra. Dopamine cells were identified electrophysiologically from whole-cell records by responses to current pulses that were typical of dopamine neurons, including hyperpolarizing 'sag' (Fig. 4c, inset). Phasic dopamine release by evoked by UV flash uncaging of caged dopamine elicited hyperpolarizations (-14.8 ± 2.0 mV amplitude, $n = 9$), which were blocked by the application of D2R antagonist, sulpiride (0.5 μM, -0.22 ± 0.54 mV, $n = 5$, $t(12) = 5.30$, $P = 0.00019$, independent samples *t*-test vs. control) consistent with the physiological actions of dopamine (Lacey *et al.*, 1987; Mercuri *et al.*, 1997). Thus, we confirmed that phasic dopamine release from caged dopamine was physiologically functional.

To test the synaptic eligibility trace hypothesis directly, we used photolytic uncaging of caged dopamine to produce phasic release of dopamine at different time points in relation to pre–post spike timing. The voltammetry measures of the timecourse of dopamine concentration changes provided an estimate of the timecourse of dopamine concentration changes shown schematically in Fig. 4d.

We then tested the eligibility trace hypothesis, by measuring the effect of dopamine applied after pre–post conjunction of activity. As shown in Fig. 4e, phasic dopamine release by uncaging 2s after pre–post conjunction of activity caused t-LTP in D1 MSNs, in the presence of nimodipine, in contrast to control conditions (no uncaging) in which no change is seen ($27.2 \pm 8.81\%$, $n = 7$, $t(12) = 2.16$, $P = 0.009$). This group comparison supports the eligibility trace hypothesis by showing that dopamine is sufficient for t-LTP when applied after a delay of 2s. While this shows that t-LTP is induced when phasic dopamine is applied 2s after pre–post pairing activity, it also raises the question whether the delay is necessary. To answer this requires examination of the range of advanced or delayed timing over which t-LTP can be induced. It also raises the question whether t-LTP would be hidden by t-LTD in the absence of nimodipine.

In order to determine the temporal requirements for dopamine in t-LTP, we systematically varied the time of dopamine release in relation to pre–post conjunction of activity. We first performed experiments in the absence of nimodipine. As shown in Fig. 4f, there was no significant difference in the changes in EPSPs between control (D1 pre–post no-nimodipine) and uncaging of dopamine at −2s, 0s or +4s (−2s, $t(12) = 0.38$, $P = 0.71$; 0s, $t(14) = 0.24$, $P = 0.82$; 4s, $t(13) = 0.05$, $P = 0.96$). However, when dopamine was uncaged at +2s, there was a significant difference from control ($t(14) = 3.38$, $P = 0.004$) and instead of t-LTD normally seen in these conditions, there was no change. Thus, dopamine uncaged 2s reverses the depression that normally occurs after pre–post conjunction of activity.

The absence of t-LTD after pre–post conjunction of activity pairs with dopamine uncaged 2s later could be due either to blocking of t-LTD by the uncaged dopamine or by co-occurrence of t-LTD and t-LTP. We therefore tested whether blocking t-LTD would block the effect of dopamine or alternatively unmask t-LTP occluded by t-LTD. To do this, we repeated the uncaging experiment in the presence of nimodipine to block t-LTD.

In the presence of nimodipine to block t-LTD, we measured t-LTP when phasic dopamine release was induced 2s after pre–post conjunction of activity, but no t-LTP when dopamine release was induced at earlier or later time points (−2 s: $-8.2 \pm 20.4\%$, $n = 5$, $t(10) = 0.48$, $P = 0.64$; 0s: $6.0 \pm 15.2\%$, $n = 4$, $t(9) = 0.58$, $P = 0.58$; +2 s: (as reported above for Fig. 4e) $27.2 \pm 8.81\%$, $n = 7$, $t(12) = 2.16$, $P = 0.009$; +4 s: $4.7 \pm 14.1\%$, $n = 4$, $t(9) = 0.46$, $P = 0.65$, Fig. 4g). These findings support the eligibility trace hypothesis by showing that the phasic dopamine signal causes t-LTP, when it occurs 2s after pre–post conjunction of activity. They also show that the eligibility trace cannot be mediated by Ca^{2+} entry through voltage-sensitive calcium channels, because t-LTP occurred when these were blocked by nimodipine.

The critical time dependence for modulation of t-LTD and t-LTP by phasic dopamine release may help to resolve some of the apparently different findings from different laboratories by considering phasic release of dopamine caused by the electrical stimulation used in the induction protocols. Although we found that APV did not block t-LTP, previous studies have reported NMDA-dependent LTP (Pawlak & Kerr, 2008; Shen *et al.*, 2008). It is possible that these forms of LTP – which are also dopamine-dependent – require NMDA receptor activation to release dopamine. To investigate the possibility that these forms of LTP are mediated by associated NMDA-dependent dopaminergic release, we measured dopamine release in response to electrical stimulation in patterns and intensities used in previous studies. Intrastriatal stimulation via theta glass electrodes, at an intensity which induced NMDA-dependent t-LTD in previous studies (Shen *et al.*, 2008), caused phasic dopamine release ([DA]peak = 138.4 ± 30.0 nM at 20 μA, $n = 10$; Fig. 5a–d). Similarly, cortical stimulation with a bipolar stimulating electrode (Fig. 5e) at an intensity reported to induce NMDA-dependent t-LTD in the presence of bicuculline (Pawlak & Kerr, 2008) also induced dopamine release ([DA]peak = 11.6 ± 9.2 nM at 0.5 mA, $n = 7$). Conversely, the NMDA receptor antagonist APV reduced dopamine release by 53.3% ([DA] peak = 5.4 ± 5.4 nM at 0.5 mA, $n = 7$, Fig. 5f), indicating that dopamine release during t-LTD induction is strongly NMDA-dependent. In contrast, no-dopamine release was detected during stimulation with the protocol used in the present experiments (cortical stimulation, 0.01–0.5 mA, 0.2 ms, no bicuculline, Fig. 5f).

Discussion

We report experimental evidence for a silent synaptic eligibility trace in the corticostriatal pathway and an estimate of its timecourse. A synaptic eligibility trace has been hypothesized to bridge a time delay between neural activity and reinforcing effects of dopamine at the cellular level. To experimentally test the eligibility trace hypothesis in the corticostriatal pathway, we first established the conditions required for t-LTD and dopamine-dependent t-LTP. Pairing of presynaptic and postsynaptic activity in sequence (pre–post) caused t-LTD in dopamine D1 receptor expressing MSNs but not in D2 MSNs. Reversing the sequence (post–pre) caused no change in D1 or D2 MSNs. A supralinear increase in spine $[Ca^{2+}]_i$ was associated with t-LTD exclusively in dendritic spines of D1 but not D2 MSNs, during pre–post but not post–pre pairing. Focussing on D1 MSNs, we found that applying a dopamine D1 receptor agonist during and after pre–post activity led to t-LTP, which was hidden by

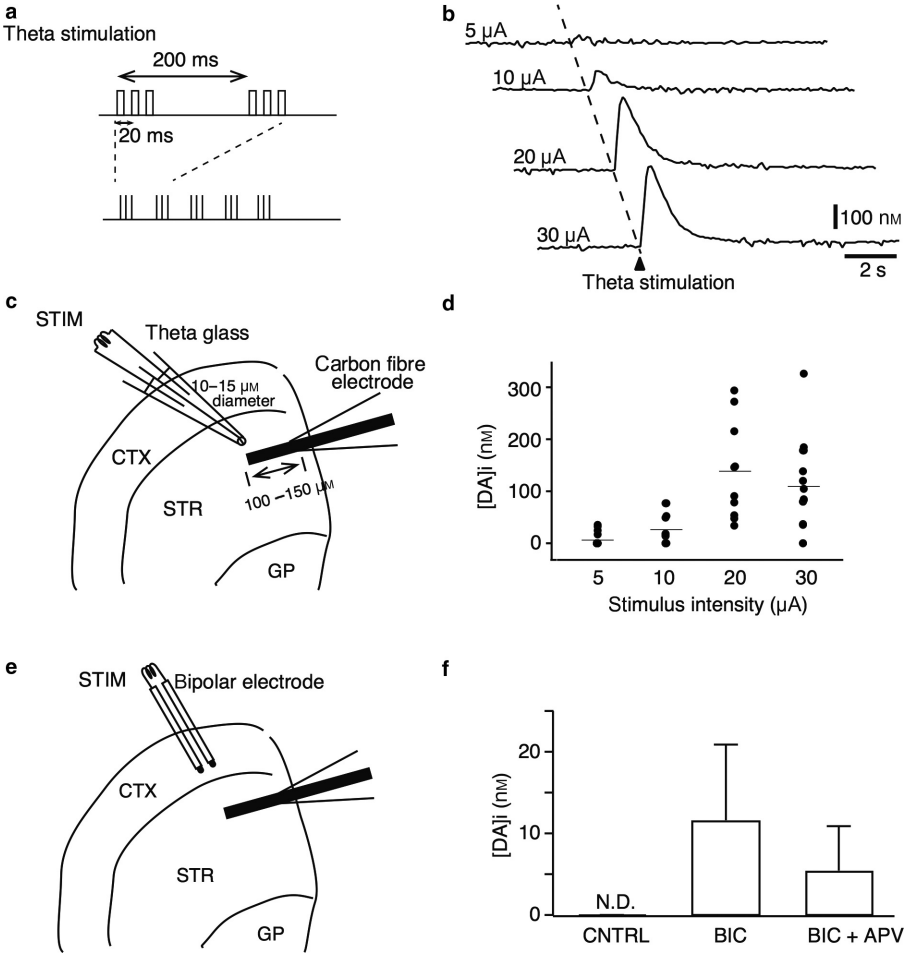

FIG. 5. Dopamine release in response to spike-timing-dependent plasticity protocols. (a) Theta stimulation protocol used to induce NMDA receptor-dependent t-LTD by Shen *et al.* (2008). (b) Voltammetric records show dopamine release evoked by theta burst stimulation at different current intensities. (c) Diagram showing intrastriatal location of theta glass stimulating electrode and carbon fibre voltammetry electrode. (d) Group average effects of theta stimulation showing dopamine release over the range of current intensities (5–30 μA). (e) Diagram showing location of cortical bipolar stimulating electrode and carbon fibre voltammetry electrode. (f) Peak dopamine concentration measured under control, bicuculline (BIC) and bicuculline plus APV conditions used by Pawlak & Kerr (2008), showing dopamine release in bicuculline is reduced by APV.

concomitant t-LTD and revealed when t-LTD was blocked by a VSCC antagonist. We then made a direct test of the eligibility trace hypothesis, using uncaging of caged dopamine to produce phasic dopamine release at different time points before and after pre–post activity. We show for the first time that phasic dopamine release occurring seconds after the cessation of presynaptic cortical and postsynaptic striatal pairing activity induced t-LTP. Dopamine release at other times did not. To the best of our knowledge, this is the first direct experimental evidence for a corticostriatal eligibility trace for reinforcement learning as proposed decades ago in neural (Miller, 1988; Wickens, 1990; Kotter & Wickens, 1995) and machine learning (Barto *et al.*, 1981, 1983, 1990) contexts, and considered essential for temporal difference learning (Suri & Schultz, 1998).

Previous work has provided promising indications of various forms of eligibility trace. These include consolidation of already-expressed synaptic plasticity based on synaptic tagging mechanisms that convert shorter forms of LTP into longer term, protein synthesis-dependent forms (Frey & Morris, 1997). The consolidation of a short-term change into a more durable form on a timescale of minutes contrasts with a silent and transient eligibility trace that is not initially expressed as LTP, but may be converted to LTP if dopamine is applied in a critical time window. Similarly, in the hippocampus, Brzosko *et al.* (2015) showed that after a prolonged period of repeated pairing, bath application of dopamine for 10–12 min during continued stimulation caused LTP. In the Brzosko *et al.* (2015) experiments, continued stimulation was necessary and dopamine application alone, after pairing, failed to induce LTP. The requirement for continued synaptic activity during the period of eligibility is inconsistent with the concept of a synaptic eligibility trace. More consistent with the eligibility trace concept, Cassenaer & Laurent (2012) reported in insects, examples of conditional modulation of t-LTD by injection of octopamine after pairing conditionally depressed contacts between cells activated in a narrow timing interval. Similarly, in rodent cerebral cortex He *et al.* (2015) reported on the Hebbian induction of silent eligibility traces, which are converted into LTP and LTD by subsequent monoamine application. In the mouse striatum, Yagishita *et al.* (2014) showed critical timing requirements for dopamine modulation of plasticity using optogenetic stimulation of mesostriatal fibres. However, in their protocol dopamine activity was induced during, but not after, the pre- and postsynaptic pairing activity, thus demonstrating critical timing requirements but not showing silent eligibility traces outlasting glutamatergic excitation. Recently, Fisher *et al.* (2017), reported in

anaesthetized rats, that a delayed sensory stimulus paired with stimulation of the substantia nigra pars compacta modulated corticostriatal spike-timing dependent plasticity in a manner consistent with an eligibility trace. However, the pairing of presynaptic and postsynaptic activity was applied at the onset of an up state, during which the barrages of afferent activity that sustain the up state continued, so that there was no silent period free of glutamatergic excitation. The current paper goes beyond previous work by demonstrating a silent and transient eligibility trace that outlasts the period of stimulation by 2s.

Establishing the existence of a synaptic eligibility trace and estimating its duration is a first step towards understanding its physical nature. Although the conditions required for eligibility and t-LTD appear similar, the present work shows that eligibility and t-LTD are two independent processes, as t-LTP and, by implication, eligibility can occur when t-LTD is blocked by a VSCC antagonist. Blocking t-LTD shifts the plasticity in the direction of t-LTP without changing the timecourse of eligibility. Thus, eligibility does not require Ca^{2+} entry via VSCCs. Also, as t-LTP is not blocked by an NMDA antagonist, Ca^{2+} entry via NMDA receptor-associated channels is not required. However, based on commonly understood mechanisms of plasticity it is reasonable to expect that Ca^{2+} entry is necessary for t-LTP. Consistent with this, we found that dopamine-dependent t-LTP was blocked by the Ca^{2+}-permeable AMPA receptor antagonist, NAS. When we directly monitored rapid changes in the fraction of Ca^{2+}-permeable AMPA receptors by measuring inward rectification properties of EPSCs interleaved with pre–post protocols (Bowie & Mayer, 1995; Kamboj *et al.*, 1995; Koh *et al.*, 1995; Plant *et al.*, 2006), we found a significant increase in inward rectification index at 2s and 4s after pre–post stimulation. While a transient increase in inward rectification after pre–post stimulation is consistent with the hypothesis that transient expression of Ca^{2+}-permeable AMPA receptors may be an important underlying component of an eligibility trace mechanism, the timecourse is more prolonged than the eligibility trace. This indicates that although there is an association with transient expression of Ca^{2+}-permeable AMPA receptors, other mechanisms must be invoked to explain the timecourse of eligibility. Previous studies have associated Ca^{2+}-permeable AMPA receptors (Shepherd, 2012) with labile memory traces that remain modifiable for a limited time interval. The involvement of Ca^{2+}-permeable AMPA receptors thus provides a promising clue to the mechanism of the eligibility trace. We speculate that the transient synaptic expression of Ca^{2+}-permeable AMPA receptors we observed may be converted into longer-lasting increase in AMPA receptor function by dopamine, acting via D1 receptors, and activation of adenylate cyclase.

Our finding that the NMDA receptor antagonist APV did not block t-LTP can be reconciled with previous studies of activity-dependent plasticity in which NMDA receptor antagonists reduced t-LTP in MSNs (Pawlak & Kerr, 2008; Shen *et al.*, 2008) if dopamine release caused by t-LTP-inducing protocols used in the previous studies is taken into account. The earliest studies of NMDA-dependent LTP in the striatum showed that Ca^{2+} entry via NMDA channels was not sufficient for LTP. Even under low magnesium conditions that increased Ca^{2+} influx via NMDA channels, LTP is blocked by dopamine D-1 receptor antagonists (Kerr & Wickens, 2001a) or dopamine depletion (Centonze *et al.*, 1999). Thus, in these earlier studies, Ca^{2+} influx via NMDA receptor-associated channels was necessary but not sufficient for LTP, raising the possibility that NMDA receptor activity may, alternatively, be related to dopamine release. In support of this interpretation, we measured dopamine release in response to electrical stimulation in patterns and intensities used in previous studies. Dopamine release occurred in both conditions that induced NMDA-dependent t-LTP in previous studies (Pawlak & Kerr, 2008; Shen *et al.*, 2008). Intrastriatal electrical stimulation via theta glass electrodes (Shen *et al.*, 2008) caused dopamine release that was reduced by the NMDA antagonist APV. Consistent with this, no NMDA-dependent t-LTP was seen in D1 neurons from dopamine-depleted slices (Shen *et al.*, 2008). Similarly, cortical stimulation at current intensities used by Pawlak & Kerr (2008) caused dopamine release in the presence of bicuculline. Thus, we propose that LTP-blocking effects of NMDA antagonists in previous studies may be mediated by their effects of NMDA antagonists on dopamine release. Variations in the relative timing of dopamine release in different spike-timing-dependent protocols may contribute to some of the differences in synaptic plasticity reported in the literature. The current findings also show that combining pre–post induction protocols with bath application of a dopamine D-1 receptor agonist is sufficient to cause t-LTP. This finding does not contradict the temporal requirements reported for phasic dopamine, because the D-1 receptor agonist was present throughout the stimulation, including the period 2s after pre–post activity. It is difficult to apply synthetic dopamine agonists by uncaging because the caged D-1 receptor agonists are not available. However, in preliminary experiments we have shown that phasic release of a D-1 agonist from a liposomal nanostructure, after a conjunction of presynaptic train stimulation and postsynaptic depolarization, can cause LTP in striatal neurons (Nakano *et al.*, 2016).

The novel neuromodulator timing-dependent plasticity we describe may play a fundamental role in learning and memory mechanisms of the corticostriatal system. The timing requirements are consistent with the delay of reinforcement gradient seen when dopamine reinforcement is given directly by intracranial stimulation, in the absence of reward-predicting cues (Black *et al.*, 1985). The duration of the silent eligibility trace we measured is also necessary and sufficient to support temporal difference learning to enable longer term predictions of reward.

Acknowledgements

Financial support from the Okinawa Institute of Science and Technology Graduate University is gratefully acknowledged. We thank Andrew Liu for technical support.

Conflict of interest statement

The authors declare no conflict of interest.

Data accessibility statement

Data files were uploaded with manuscript.

Author contributions

TS and MS were involved in conceptualization, experimental design, data collection, statistical analysis and writing of the manuscript. SW performed statistical analysis, data collection, and reviewed the manuscript. JW was involved in conceptualization, experimental design, statistical analysis, writing of the manuscript and supervision.

Abbreviations

AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; NMDA, *N*-methyl-D-aspartate; MSN, Medium-sized spiny neuron; FSCV, Fast-scan cyclic voltammetry; EPSP, Excitatory postsynaptic potential; uEPSP, Uncaging excitatory postsynaptic potential; EPSC, Excitatory postsynaptic

current; EGFP, Recombinant Enhanced Green Fluorescent Protein; APV, (2*R*)-amino-5-phosphonopentanoate; NAS, 1-naphthylacetyl spermine; VSCC, Voltage-sensitive calcium channel; DA, Dopamine; t-LTD, Spike-timing-dependent long-term depression; t-LTP, Spike-timing-dependent long-term potentiation.

References

Barto, A.G., Sutton, R.S. & Brouwer, P.S. (1981) Associative search network: A reinforcement learning associative memory. *Biol. Cybern.*, **40**, 201–211.

Barto, A.G., Sutton, R.S. & Anderson, C.W. (1983) Neuronlike elements that can solve difficult learning control problems. *IEEE Trans. Syst. Man Cyber.*, **15**, 835–846.

Barto, A.G., Sutton, R.S. & Watkins, C.J.C.H. (1990). Learning and sequential decision making. In Gabriel, M. & Moore, J.W. (Eds), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. MIT Press, Cambridge, MA, pp. 539–602.

Black, J., Belluzzi, J.D. & Stein, L. (1985) Reinforcement delay of one-second severely impairs acquisition of brain self-stimulation. *Brain Res.*, **359**, 113–119.

Bowie, D. & Mayer, M.L. (1995) Inward rectification of both AMPA and kainate subtype glutamate receptors generated by polyamine-mediated ion channel block. *Neuron*, **15**, 453–462.

Brzosko, Z., Schultz, W. & Paulsen, O. (2015) Retroactive modulation of spike timing-dependent plasticity by dopamine. *eLife*, **4**, e09685.

Calabresi, P., Maj, R., Pisani, A., Mercuri, N.B. & Bernardi, G. (1992) Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J. Neurosci.*, **12**, 4224–4233.

Cassenaer, S. & Laurent, G. (2012) Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, **482**, 47–52.

Centonze, D., Gubellini, P., Picconi, B., Calabresi, P., Giacomini, P. & Bernardi, G. (1999) Unilateral dopamine denervation blocks corticostriatal LTP. *J. Neurophysiol.*, **82**, 3575–3579.

Cheer, J.F., Aragona, B.J., Heien, M.L., Seipel, A.T., Carelli, R.M. & Wightman, R.M. (2007) Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. *Neuron*, **54**, 237–244.

Fino, E., Deniau, J.M. & Venance, L. (2008) Cell-specific spike-timing-dependent plasticity in GABAergic and cholinergic interneurons in corticostriatal rat brain slices. *J. Physiol.*, **586**, 265–282.

Fiorillo, C.D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1898–1902.

Fisher, S.D., Robertson, P.B., Black, M.J., Redgrave, P., Sagar, M.A., Abraham, W.C. & Reynolds, J.N.J. (2017) Reinforcement determines the timing dependence of corticostriatal synaptic plasticity *in vivo*. *Nat. Commun.* **8**, 334.

Frey, U. & Morris, R.G.M. (1997) Synaptic tagging and long-term potentiation. *Nature*, **385**, 533–536.

Gertler, T.S., Chan, C.S. & Surmeier, D.J. (2008) Dichotomous anatomical properties of adult striatal medium spiny neurons. *J. Neurosci.*, **28**, 10814–10824.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A. *et al.* (2003) A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature*, **425**, 917–925.

He, K., Huertas, M., Hong, S.Z., Tie, X., Hell, J.W., Shouval, H. & Kirkwood, A. (2015) Distinct eligibility traces for LTP and LTD in cortical synapses. *Neuron*, **88**, 528–538.

Izhikevich, E.M. (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex*, **17**, 2443–2452.

Kamboj, S.K., Swanson, G.T. & Cull-Candy, S.G. (1995) Intracellular spermine confers rectification on rat calcium-permeable AMPA and kainate receptors. *J. Physiol.*, **486**(Pt 2), 297–303.

Kerr, J.N. & Wickens, J.R. (2001) Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *J. Neurophysiol.*, **85**, 117–124.

Koh, D.S., Burnashev, N. & Jonas, P. (1995) Block of native Ca(2+)-permeable AMPA receptors in rat brain by intracellular polyamines generates double rectification. *J. Physiol.*, **486**(Pt 2), 305–312.

Kotter, R. & Wickens, J.R. (1995) Interactions of glutamate and dopamine in a computational model of the striatum. *J. Comput. Neurosci.*, **2**, 195–214.

Lacey, M.G., Mercuri, N.B. & North, R.A. (1987) Dopamine acts on D2 receptors to increase potassium conductance in neurones of the rat substantia nigra zona compacta. *J. Physiol.*, **392**, 397–416.

Lee, T.H., Gee, K.R., Ellinwood, E.H. & Seidler, F.J. (1996) Combining 'caged dopamine' photolysis with fast-scan cyclic voltammetry to assess dopamine clearance and release autoinhibition in vitro. *J. Neurosci. Methods*, **67**, 221–231.

Lovinger, D.M., Tyler, E.C. & Marritt, A. (1993) Short- and long-term depression in the rat neostriatum. *J. Neurophysiol.*, **70**, 1937–1949.

Mercuri, N.B., Saiardi, A., Bonci, A., Picetti, R., Calabresi, P., Bernardi, G. & Borrelli, E. (1997) Loss of autoreceptor function in dopaminergic neurons from dopamine D2 receptor deficient mice. *Neuroscience*, **79**, 323–327.

Miller, R. (1988). Cortico-striatal and cortico-limbic circuits: A two tiered model of learning and memory function. In Markowitsch, H. (Ed), *Information Processing by the Brain: Views and Hypotheses from a Cognitive-Physiological Perspective*. Hans Huber Press, Bern, pp. 179–198.

Nakano, T., Mackay, S.M., Wui Tan, E., Dani, K.M. & Wickens, J. (2016) Interfacing with neural activity via femtosecond laser stimulation of drug-encapsulating liposomal nanostructures. *eNeuro*, **3**, ENEURO.0107-16.2016.

Paille, V., Fino, E., Du, K., Morera-Herreras, T., Perez, S., Kotaleski, J.H. & Venance, L. (2013) GABAergic circuits control spike-timing-dependent plasticity. *J. Neurosci.*, **33**, 9353–9363.

Pawlak, V. & Kerr, J.N. (2008) Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.*, **28**, 2435–2446.

Plant, K., Pelkey, K.A., Bortolotto, Z.A., Morita, D., Terashima, A., McBain, C.J., Collingridge, G.L. & Isaac, J.T. (2006) Transient incorporation of native GluR2-lacking AMPA receptors during hippocampal long-term potentiation. *Nat. Neurosci.*, **9**, 602–604.

Reynolds, J.N., Hyland, B.I. & Wickens, J.R. (2001) A cellular mechanism of reward-related learning. *Nature*, **413**, 67–70.

Schultz, W., Preuschoff, K., Camerer, C., Hsu, M., Fiorillo, C.D., Tobler, P.N. & Bossaerts, P. (2008) Explicit neural signals reflecting reward uncertainty. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **363**, 3801–3811.

Shen, W., Flajolet, M., Greengard, P. & Surmeier, D.J. (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*, **321**, 848–851.

Shepherd, J.D. (2012) Memory, plasticity and sleep – A role for calcium permeable AMPA receptors? *Front. Mol. Neurosci.*, **5**, 49.

Shindou, T., Ochi-Shindou, M. & Wickens, J.R. (2011) A Ca(2+) threshold for induction of spike-timing-dependent depression in the mouse striatum. *J. Neurosci.*, **31**, 13015–13022.

Suri, R.E. & Schultz, W. (1998) Learning of sequential movements by neural network model with dopamine- like reinforcement signal. *Exp. Brain Res.*, **121**, 350–354.

Thivierge, J.P., Rivest, F. & Monchi, O. (2007) Spiking neurons, dopamine, and plasticity: timing is everything, but concentration also matters. *Synapse*, **61**, 375–390.

Wickens, J.R. (1990) Striatal dopamine in motor activation and reward-mediated learning. Steps towards a unifying model. *J. Neural. Transm.*, **80**, 9–31.

Wickens, J.R., Begg, A.J. & Arbuthnott, G.W. (1996) Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience*, **70**, 1–5.

Wu, Q., Reith, M.E., Wightman, R.M., Kawagoe, K.T. & Garris, P.A. (2001) Determination of release and uptake parameters from electrically evoked dopamine dynamics measured by real-time voltammetry. *J. Neurosci. Methods*, **112**, 119–133.

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C., Urakubo, H., Ishii, S. & Kasai, H. (2014) A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science*, **345**, 1616–1620.