Article:
Hunger, L., Kumar, A. and Schmidt, R. orcid.org/0000-0002-2474-3744 (2020) Abundance compensates kinetics : similar effect of dopamine signals on D1 and D2 receptor populations. Journal of Neuroscience, 40 (14). pp. 2868-2881. ISSN 0270-6474

https://doi.org/10.1523/JNEUROSCI.1951-19.2019

Title: Abundance compensates kinetics: Similar effect of dopamine signals on D1 and D2 receptor populations

Abbreviated title: Dopamine receptor abundance compensates kinetics

Lars Hunger[1], Arvind Kumar[2], Robert Schmidt[1]

[1]Department of Psychology, University of Sheffield, UK

[2]Computational Science and Technology, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden

Corresponding author:

Lars Hunger

Department of Psychology, University of Sheffield, UK

pc4xlh@sheffield.ac.uk, Lars.Hunger.314@gmail.com

Number of pages: 43

Number of figures: 11

Number of tables: 1

Number of words for Abstract: 194

Number of words for Significance statement: 115

Number of words for Introduction: 655

Number of words for Discussion: 1240

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgements: We thank Joshua Berke, Paul Overton, Alejandro Jimenez, Mohammadreza Mohagheghi Nejad and Amin Mirzaei for helpful discussions. This work was supported by the University of Sheffield and its high performance computing resources, by funding from the EU H2020 Programme as part of the Human Brain Project (HBP-SGA1, 720270; HBP-SGA2, 785907), and the BrainLinks-BrainTools Cluster of Excellence funded by the German Research Foundation (DFG, grant number EXC 1086), and the state of Baden-Wuerttemberg through bwHPC.

Abstract

The neuromodulator dopamine plays a key role in motivation, reward-related learning and normal motor function. The different affinity of striatal D1 and D2 dopamine receptor types has been argued to constrain the D1 and D2 signalling pathways to phasic and tonic dopamine signals, respectively. However, this view assumes that dopamine receptor kinetics are instantaneous so that the time courses of changes in dopamine concentration and changes in receptor occupation are basically identical. Here we developed a neurochemical model of dopamine receptor binding taking into account the different kinetics and abundance of D1 and D2 receptors in the striatum. Testing a large range of behaviorally-relevant dopamine signals, we found that the D1 and D2 dopamine receptor populations responded very similarly to tonic and phasic dopamine signals. Furthermore, due to slow unbinding rates, both receptor populations integrated dopamine signals over a timescale of minutes. Our model provides a description of how physiological dopamine signals translate into changes in dopamine receptor occupation in the striatum, and explains why dopamine ramps are an effective signal to occupy dopamine receptors. Overall, our model points to the importance of taking into account receptor kinetics for functional considerations of dopamine signalling.

Significance statement

Current models of basal ganglia function are often based on a distinction of two types of dopamine receptors, D1 and D2, with low and high affinity, respectively. Thereby, phasic dopamine signals are believed to mostly affect striatal neurons with D1 receptors, and tonic dopamine signals are believed to mostly affect striatal neurons with D2 receptors. This view does not take into account the rates for the binding and unbinding of dopamine to D1 and D2 receptors. By incorporating these kinetics into a computational model we show that D1 and D2 receptors both respond to phasic and tonic dopamine signals. This has implications for the processing of reward-related and motivational signals in the basal ganglia.

Introduction

The neuromodulator dopamine (DA) plays a key role in motivation, reward-related learning and normal motor function. Many aspects of DA function are mediated by its effects on the excitability (Day et al., 2008) and strength of cortico-striatal inputs (Reynolds et al., 2001) in the context of motor control (Syed et al., 2016), action-selection (Redgrave et al., 2010), reinforcement learning (Schultz, 2007), and addiction (Everitt and Robbins, 2005). The striatal DA concentration ([DA]) can change over multiple timescales (Schultz, 2007). Fast increases in [DA] lasting for $\approx 1-3s$ result from phasic bursts in DA neurons (Roitman et al., 2004), which signal reward-related information (Schultz, 2007; Grace et al., 2007). Slightly slower [DA] ramps occur as animals approach goal locations (Howe et al., 2013) or perform reinforcement learning tasks (Hamid et al., 2016). Finally, tonic firing of DA neurons may control the baseline [DA] and change on a timescale of minutes or longer (Grace et al., 2007). However, whether e.g. learning and motivation are mediated by different timescales of DA cell firing (Niv et al., 2007) has recently been challenged (Berke, 2018; Mohebi et al., 2019). The issue of DA signalling time scales is important because D1 and D2 DA receptors may react to different timescales of the DA signal due to their different affinities for DA.

Based on the different DA affinities of D1 and D2 receptors (D1R and D2R), it is often assumed that striatal medium spiny neurons (MSNs) respond differently to tonic and phasic DA signals, depending on which DA receptor type they predominantly express (Dreyer et al., 2010; Surmeier et al., 2007; Grace et al., 2007; Schultz, 2007; Frank and O'Reilly, 2006). According to this "affinity-based" model, the low affinity D1Rs (high dissociation constant $K_D^{D1} = 1.6\mu M$; Richfield et al., 1989) cannot detect tonic changes in [DA] because the fraction of occupied D1Rs is too small ($\approx 1\%$) at a baseline [DA] of $20nM$ and does not change much during tonic, low amplitude [DA] changes. However, D1Rs can detect phasic, high amplitude [DA] increases because they only saturate at a very high [DA]. By contrast, D2Rs have a high affinity (low dissociation constant $K_D^{D2} = 25nM$; Richfield et al., 1989) leading to $\approx 40\%$ of D2Rs being occupied at a baseline [DA] of $20nM$. Due to their high affinity, D2Rs can detect low amplitude, tonic [DA] increases/decreases. However, D2Rs saturate at relatively low $[DA] > 2 \cdot K_D^{D2}$, and therefore cannot detect high amplitude, phasic [DA] increases. This suggests that D1 and D2 type MSNs respond differently to phasic and tonic [DA] changes because of the different affinities of D1Rs and D2Rs (Schultz, 2007).

However, this model neglects other factors relevant for receptor occupation and is incompatible with recent findings that D2R expressing MSNs can detect phasic [DA] changes (Marcott et al., 2014; Yapo et al., 2017).

The affinity-based model assumes that the reaction equilibrium is reached instantaneously, whereby the affinity can be used to approximate the fraction of receptors bound to DA. However, this assumption holds only if the receptor kinetics are fast compared to the timescale of the DA signal, which is typically not the case. For instance, D1Rs and D2Rs unbind from DA with a half-life time of $t_{1/2} \approx 80s$ (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Nishikori et al., 1980), much longer than phasic signals of a few seconds (Robinson et al., 2001; Schultz, 2007; Hamid et al., 2016). Moreover, the fraction of bound receptors might be a misleading measure for the effect of DA signals, since the abundances of D1R and D2R in the striatum are quite different. Abundance here refers to the total number of D1 or D2 receptors available to bind to DA within a volume of extracellular space.

To investigate the role of receptor kinetics and abundances for DA signalling in the striatum, we developed a neurochemical model incorporating kinetics and abundances of D1Rs and D2Rs, and re-evaluated current views on DA signalling in the striatum.

Methods and Materials

Code Accessibility

All models were implemented in Python. The models and all scripts used to generate the data and figures can be accessed here:
`https://bitbucket.org/Narur/abundance_kinetics/src/` .

Kinetics model

In the affinity-based model the receptor kinetics are instantaneous, so that the fraction of occupied D1 and D2 receptors (f_{D1} and f_{D2}) can be calculated directly from the concentration of free DA in the extracellular space, $[DA]$, and the dissociation constant K_D (see e.g.

Copeland 2004):

$$f = \frac{[DA]}{K_D + [DA]}. \tag{1}$$

However, the dissociation constant is an equilibrium constant, so it should only be used for calculating the receptor occupancy when the duration of the DA signal is longer than the time needed to reach the equilibrium. As this is typically not the case for phasic DA signals, since the half-life time of receptors is longer (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Nishikori et al., 1980) than the timeframe of phasic signaling (Roitman et al., 2004), we developed a model which incorporates slow kinetics.

When DA and one of its receptors are both present in a solution they constantly bind and unbind. During the binding process a receptor ligand complex (here called DA−D1 or DA−D2) is formed (see e.g. Copeland 2004). We refer to the receptor ligand complex as an occupied DA receptor. Below we provide the model equations for D1 receptors, but the same equations apply for D2 receptors (with different kinetic parameters). In a solution binding occurs when receptor and ligand meet due to diffusion, with high enough energy and a suitable orientation, described as:

$$\mathrm{DA} + \mathrm{D1} \xrightarrow{\mathrm{k_{on}}} \mathrm{DA{-}D1}. \tag{2}$$

Accordingly, unbinding of the complex is denoted as:

$$\mathrm{DA{-}D1} \xrightarrow{\mathrm{k_{off}}} \mathrm{DA} + \mathrm{D1}. \tag{3}$$

The kinetics of this binding and unbinding, treated here as first-order reactions, are governed by the rate constants k_{on} and k_{off} that are specific for a receptor ligand pair and temperature dependent. Since both processes are happening simultaneously we can write this as:

$$\mathrm{DA} + \mathrm{D1} \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}}}{\rightleftharpoons}} \mathrm{DA{-}D1}. \tag{4}$$

The rate at which the receptor is occupied depends on [DA], the concentration of free receptor [D1] and the binding rate constant k_{on}:

$$\frac{d[\mathrm{DA-D1}]^{+}}{dt} = k_{\mathrm{on}} \cdot [\mathrm{DA}] \cdot [\mathrm{D1}]. \tag{5}$$

The rate at which the receptor-ligand complex unbinds is given by the concentration of the complex $[\mathrm{DA-D1}]$ and the unbinding rate constant k_{off}:

$$\frac{d[\mathrm{DA-D1}]^{-}}{dt} = -k_{\mathrm{off}} \cdot [\mathrm{DA-D1}]. \tag{6}$$

The equilibrium is reached when the binding and unbinding rates are equal, so by combining Eq. 5 and Eq. 6 we obtain:

$$k_{\mathrm{on}} \cdot [\mathrm{DA}] \cdot [\mathrm{D1}] = k_{\mathrm{off}} \cdot [\mathrm{DA} - \mathrm{D1}]. \tag{7}$$

At the equilibrium the dissociation constant K_D is defined as:

$$K_D = \frac{[DA] \cdot [D1]}{[DA - D1]} = \frac{k_{off}}{k_{on}}. \tag{8}$$

When half of the receptors are occupied, i.e. $[DA - D1] = [D1]$, Eq. 8 simplifies to $K_D = [DA]$. So at equilibrium, K_D is the ligand concentration at which half of the receptors are occupied.

Importantly, for fast changes in [DA] (i.e. over seconds) it takes some time until the changed binding (Eq. 5) and unbinding rates (Eq. 6) are balanced, so the new equilibrium will not be reached instantly. The timescale in which equilibrium is reached can be estimated from the half-life time of the bound receptor. The half-life time assumes an exponential decay process as described in Eq. 6 and is the time required so that half of the currently bound receptors unbind. Although the $t_{1/2}$ estimates have been published, it is also possible to calculate it from experimental estimates of k_{off} by using $t_{1/2} = ln(2)/k_{off}$ (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Nishikori et al., 1980). Signal duration should be of the same order of magnitude or longer than the half-life time in order for the affinity-based model with instant kinetics to be applicable.

We calculated the time course of occupied receptor after an abrupt change in [DA] by integrating the rate equation, given by the sum of Eq. 5 and Eq. 6:

$$\frac{d[DA - D1]}{dt} = k_{on}[DA][D1] - k_{off}[DA - D1]. \tag{9}$$

To integrate Eq. 9 we substitute

$$[D1] = [D1^{tot}] - [DA - D1] \tag{10}$$

where $[D1^{tot}]$ is the total amount of D1 receptor (bound and unbound to DA) on the cell membranes available for binding to extracellular DA.

To model the effect of phasic changes in [DA] we choose the initial receptor occupancy $[DA - D1](t = 0) = [DA - D1]^0$ and the receptor occupancy for the new equilibrium at time

infinity $[DA-D1](t=\infty)=[DA-D1]^{\infty}$ as the boundary conditions. With these boundary conditions we get an analytic expression for the time evolution of the receptor occupancy under the assumption that binding to the receptor does not significantly change the free [DA]:

$$[DA-D1](t)=([DA-D1]^{0}-[DA-D1]^{\infty})\cdot e^{-(k_{on}[DA]+k_{off})t}+[DA-D1]^{\infty}. \quad (11)$$

For arbitrary DA time courses we solve Eq. 9 for each receptor type numerically employing a 4th order Runge Kutta solver with a 1 ms time resolution.

We did not take into account the change in [DA] caused by the binding and unbinding to the receptors since the rates at which DA is removed from the system by binding to the receptors is much slower than the rate of DA being removed from the system by uptake through DA transporters. The rate at which DA binds to the receptors is:

$$\frac{d[DA-D1]}{dt}+\frac{d[DA-D2]}{dt}=k_{on}^{D1}[DA][D1]+k_{on}^{D2}[DA][D2]=-\frac{d[DA]}{dt}. \quad (12)$$

This equation (Eq. 12) relates the removal of DA to the binding of DA to its receptors. To estimate the binding of DA to its receptors we use the parameters of the DA step example (**Fig. 1**). In this example there is an instantaneous DA increase from the baseline value $[DA]=20\,nM$ to $[DA]=1\,\mu M$. At the time of the step up, the D1 and D2 occupancy is given as $[DA-D1]\approx 20.0\,nM$ and $[DA-D2]\approx 40\,nM$ (the equilibrium values for baseline DA). With that the free receptor concentration is $[D1]=[D1]^{tot}-[D1-DA]\approx 1600.0\,nM$ and $[D2]=[D2]^{tot}-[D2-DA]\approx 40.0\,nM$. The receptor parameters $k_{on}^{D1}=5.2\cdot 10^{-6}nM^{-1}s^{-1}$, $k_{on}^{D2}=3.3\cdot 10^{-4}nM^{-1}s^{-1}$, and the receptor abundances $[D1]^{tot}$ and $[D2]^{tot}$ are derived in the receptor parameters section below. For the parameters from this example, the rate of DA removal through binding to the receptors is given by:

$$\frac{[DA]}{dt}^{binding}=-23.6\,nM/s. \quad (13)$$

However, the DA removal rate by Michaelis-Menten uptake through the DA transporters at

this concentration would be:

$$\frac{[DA]^{uptake}}{dt} = V_{max}\frac{[DA]}{[DA]+K_m} \tag{14}$$

$$= -4.0\frac{\mu M}{s}\cdot\frac{1\mu M}{1\mu M+0.21\mu M} \tag{15}$$

$$= -3.3\frac{\mu M}{s}. \tag{16}$$

$V_{max} = -4.0\frac{\mu M}{s}$ is the maximal uptake rate, and $K_m = 0.21\mu M$ the Michaelis-Menten constant describing the $[DA]$ concentration at which uptake is at half the maximum rate (Bergstrom and Garris, 2003). As $\left|\frac{[DA]^{uptake}}{dt}\right| >> \left|\frac{[DA]^{binding}}{dt}\right|$, the DA dynamics are dominated by the uptake process and not by binding to the receptors. Therefore, we neglected the receptor-ligand binding for the DA dynamics in our model. However, for faster DA receptors this effect would become more important.

Receptor parameters

The DA receptor abundances, i.e. the total concentration of the D1 and D2 receptors on the membrane ($[D1]^{tot}$ and $[D2]^{tot}$) that can bind to DA in the extracellular space of the striatum are important model parameters. Our estimate of $[D1]^{tot}$ and $[D2]^{tot}$ is based on radioligand binding studies in the rat rostral striatum (Richfield et al., 1989, 1987). We use the following equation, in which X is a placeholder for the respective receptor type, to calculate these concentrations.

$$[DX]^{tot} = [DX]^m \cdot \frac{\epsilon \cdot f^m_{DX}}{\alpha\rho_b} \tag{17}$$

The experimental measurements provide us with the number of receptors per unit of protein weight $[D1]^m$ and $[D2]^m$. To transform these measurements into molar concentrations for our simulations, we multiply by the protein content of the wet weight of the rat caudate nucleus ϵ, which is around 12% (Banay-Schwartz et al., 1992). This leaves us with the amount of protein per g of wet weight of the rat brain. Next we divide by the average density of a rat brain which is $\rho_b = 1.05g/ml$ (DiResta et al., 1990) to find the amount of receptors per unit of volume of the rat striatum. Finally, we divide by the volume fraction α, the fraction of the brain volume that is taken up by the extracellular space in the rat brain, to obtain the receptor concentration of the receptor in the extracellular medium. The procedure ends here for the

D1 receptors since there is no evidence that D1 receptors are internalized in the baseline state (Prou et al., 2001; Nishikori et al., 1980). However, a large fraction of the D2 receptors is retained in the endoplasmatic reticulum of the neuron (Prou et al., 2001; Nishikori et al., 1980), reducing the amount of receptors that contribute to the concentration of receptors in the extracellular medium by f^m, the fraction of receptors protruding into the extracellular medium. Thus, we define the receptor abundances $[DX]^{tot}$, the total number of receptors per unit of volume in the extracellular medium for both receptor types. It is useful to give this quantity as a concentration (nM), so that it can be easily used for calculating binding rates and equilibria as described above.

Nishikori et al. (1980) also give estimates corresponding to $[D1]^m$ and $[D2]^m$. Their measurements give $[D1]^m = 11.9\,pmol/mg\,(protein)$ and $[D2]^m = 0.16\,pmol/mg\,(protein)$. Because these measurements are already for the cellular membrane (i.e. they have already been implicitly multiplied by $f^{membrane}$), we used **Eq. 17** to obtain $[D1]^{tot} \approx 6400\,nM$ and $[D2]^{tot} \approx 100\,nM$. Note that estimates of $[D1]^m$ may differ by a factor three among in the nucleus accumbes of rats, cats and monkeys (Richfield et al., 1987). Here, we used the values corresponding to rats (Richfield et al., 1987, 1989) instead of canine (Nishikori et al., 1980). Despite the differences across species, the receptor abundances derived from Nishikori et al. (1980) indicate that, at baseline $[DA]$, the $[D1 - DA]$ would still be of the same order of magnitude as $[D2 - DA]$. However, with these values $[D1 - DA]$ would be roughly twice $[D2 - DA]$ (as opposed to half in our model, see **Fig. 1**).

In addition to the receptor concentration, the kinetic constants of the receptors are key parameters in our slow kinetics model. In an equilibrium measurement in the canine caudate nucleus the dissociation constant of low affinity DA binding sites, corresponding to D1 receptors (Maeno, 1982), has been measured as $K_D = 1.6\mu M$ (Sano et al., 1979). However, when calculating K_D (using Eq. 8) from the measured kinetic constants (Sano et al., 1979) the value is $K_D^{D1} = 2.6\mu M$. To be more easily comparable to other simulation works (Dreyer et al., 2010) and direct measurements (Richfield et al., 1989; Sano et al., 1979) we choose $K_D^{D1} = 1.6\mu M$ in our simulations. For this purpose we modified both the $k_{on}^{D1} = 0.00025 min^{-1} nM^{-1}$ and $k_{off}^{D1} = 0.64 min^{-1}$ rate measured (Sano et al., 1979) by $\approx 25\%$, making $k_{on}^{D1} = 0.0003125 min^{-1} nM^{-1}$ slightly faster and $k_{off}^{D1} = 0.5 min^{-1}$ slightly slower, so that the resulting $K_D^{D1} = 1.6\mu M$. The kinetic constants have been measured at

30^oC and are temperature dependent. In biological reactions a temperature change of 10^oC is usually associated with a change in reaction rate around a factor of 2-3 (Reyes et al., 2008). However, the conclusions of this paper do not change for an increase in reaction rates by a factor of $2-3$ (see **Fig. 9**). It should also be noted that the measurements of the commonly referenced K_D (Richfield et al., 1989) have been performed at room temperature.

The kinetic constants for the D2 receptors were obtained from measurements at 37^oC of high affinity DA binding sites (Burt et al., 1976), which correspond to the D2 receptor (Maeno, 1982). The values are $k_{on}^{D2} = 0.02min^{-1}nM^{-1}$ and $k_{off}^{D2} = 0.5min^{-1}$, which yields $K_D^{D2} = 25nM$, in line with the values measured in (Richfield et al., 1989). As the measured off-rate of the D1 and D2 receptors $k_{off}^{D1} = 0.64min^{-1}$ and $k_{off}^{D2} = 0.5min^{-1}$ is quite similar and we modify the measured values slightly in our model (see above), the difference in $K_D^{D2} = 25nM$ and $K_D^{D1} = 1.6\mu M$ is largely due to differences in the on-rate of the receptors. This is important because the absolute rate of receptor occupancy depends linearly not only on the on-rate, but also on the receptor concentration (see Eq. 5), which means that a slower on-rate could be compensated for by a higher number of receptors.

The parameters used in the simulations are summarized in Tab. 1.

Dopamine signals

In our model we assumed a baseline [DA] of $[DA]^{tonic} = 20\ nM$ (Dreyer et al., 2010; Dreyer, 2014; Venton et al., 2003; Suaud-Chagny et al., 1992; Borland et al., 2005; Justice Jr, 1993; Atcherley et al., 2015). We modelled changes in [DA] to mimic DA signals observed in experimental studies. We use three types of single pulse DA signals: (long-)bursts, burst-pauses, and ramps.

The (long-)burst signal mimics the effect of a phasic burst in the activity of DA neurons in the SNc, e.g. in response to reward-predicting cues (Pan et al., 2005). The model burst signal consists of a rapid linear [DA] increase (with an amplitude $\Delta[DA]$ and rise time t_{rise}) and a subsequent return to baseline. The return to baseline is governed by Michaelis Menten kinetics with appropriate parameters for the dorsal striatum $V_{max} = 4.0\ \mu Ms^{-1}$ and $K_m = 0.21\ \mu M$ (Bergstrom and Garris, 2003) and the nucleus accumbens $V_{max} = 1.5\ \mu Ms^{-1}$ (Dreyer and

Hounsgaard, 2013). In our model the removal of DA is assumed to happen without further DA influx into the system (baseline firing resumes when [DA] has returned to its baseline value). Unless stated otherwise, the long-burst signals are used with a $\Delta[DA] = 200\ nM$ and a rise time of $t_{rise} = 0.2\ s$ at $V_{max} = 1.5\ \mu M s^{-1}$, similar to biologically realistic transient signals (Cheer et al., 2007; Robinson et al., 2001; Day et al., 2007).

The burst-pause signal has two components, an initial short, small amplitude burst ($\Delta[DA] = 100\ nM$, $t_{rise} = 0.1\ s$), with the corresponding [DA] returning then to baseline (as for the long burst above). However, there is a second component in the DA signal, in which [DA] falls below baseline, simulating the effect of a pause in DA neuron firing. The length of this firing pause is characterized by the parameter t_{pause}. We simulated this type of burst-pause [DA] signal to investigate how e.g. a two-component response of the DA reward prediction error (Schultz, 2016) would affect the DA receptor occupation. In this case the model input [DA] time course is based on DA cell firing patterns consisting of a brief burst followed by a pause in activity (Pan et al., 2008; Schultz, 2016).

The ramp DA signal is characterized by the same parameters as the burst pattern, but with a longer t_{rise}, and a smaller $\Delta[DA]$ (parameter settings provided in each simulation).

For the simulations comparing the area under the curve of the input DA signal with the resulting receptor occupancy (**Fig. 5**) we used the burst, burst-pause, and ramp signals described above with a range of parameter settings. For the burst DA signal we used amplitudes $\Delta[DA]^{max}$ of 50, 100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, and 1000 nM. For the ramping DA signals we used rise times t_{rise} of 0.2, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, and 7.0 s, For the burst-pause DA signal we used different values for V_{max} of 1.0, 1.5, 2.0, 2.5, 3.0, 2.5, and 4.0 $\mu M s^{-1}$.

Behavioural task simulation

To determine whether DA receptor occupancy can integrate reward signals over minutes, we simulated experiments consisting of a sequence of 50 trials. In each sequence the reward probability was fixed. The trials contained either a (long-)burst DA signal (mimicking a reward) or a burst-pause DA signal (mimicking no reward) at the beginning of the trial according to the

reward probability of the sequence. The inter-trial interval was $15 \pm 5s$ (**Fig. 8**). We choose this highly simplistic scenario to mimic DA signals in a behavioural task in which the animal receives unpredictable rewards with a given reward probability. Due to the unpredictable nature of the reward, we assumed that here the DA pulse amplitude is not affected by the reward probability, as e.g. DA cell recordings during unsignalled reward presentations on a similar time scale have been used to obtain strong DA cell responses (Fiorillo et al., 2008).

However, here the specifics of the task are not relevant as our model addresses the integration of the DA receptor occupancy over time. Although we chose to use the burst-pause type signal as shown in **Fig. 2a** as a non-rewarding event, the difference to a non-signal are minimal after the end of the pause (**Figs. 3 and 4**). Each sequence started from a baseline receptor occupancy, assuming a break between sequences long enough for the receptors to return to baseline occupancy (around 5 minutes). For the simulations shown in **Fig. 4** all trials started exactly 15 s apart.

While for the simulations shown in **Fig. 4** the sequence of DA signals was fixed, we also simulated a behavioural task with stochastic rewards (**Fig. 8**). There we simulated reward probabilities from 0% to 100% in 10% steps. For each reward probability we ran 500 sequences, and calculated the mean receptor occupancy over time (single realizations shown in **Fig. 8a, c**). To investigate whether the receptor occupancy distinguished between different reward probabilities we applied a simple classifier to the receptor occupancy time course.

The classifier was used to compare two different reward probabilities at a time. At each time point during the simulated experiment it was applied to a pair of receptor occupancies, e.g. one belonging to a 70% and one to a 30% reward probability sequence. For each sequence the classifier assigned the current receptor occupancy to the higher or lower reward probability depending on which reward probabilities' mean receptor occupancy (over 500 sequences) was closer to the current receptor occupancy.

As we knew the underlying reward probability of each sequence we were able to calculate the true and false positive rates for each time point in our set of 500 sequences for both the D1R and D2R (**Fig. 8e**). For each individual comparison of two reward probability sequences, at each time point the classifier could make a correct or incorrect classification (denoted by a '1' or '0', respectively). The true and false positive rates were then obtained by averaging

these sequences of zeros and ones at each time point over 500 different realizations of the two reward probability sequences. For example, **Fig. 8e** shows that at a time of 400s in approximately 94% of the 1000 (500 for each probability) studied sequences the underlying reward probability was correctly identified by the classifier (i.e. 30% as 30%, and 70% as 70%) based on the D1 receptor occupation. The classification was done by calculating the distance of the instantaneous D1 receptor occupancy (i.e. in this case at 400s) of a given individual sequence to the mean (over 500 sequences) receptor occupation of the 30% and 70% cases and then choosing the closer one. From the time resolved true and false positive rates we calculated the time averaged true and false positive rates, for all pairs of probabilities, (**Fig. 8b, d**) and the time averaged accuracy (**Fig. 8f**) using all time points between 200 and 800s within a sequence to avoid the effect of the initial "swing-in" and post-sequence DA levels returning to baseline.

Results

Before investigating the role of the receptor kinetics in response to different DA signals, we started by establishing the receptor binding at baseline [DA], taking into account the different abundances of D1 and D2 receptors in the striatum. For a stable baseline [DA] the resulting receptor occupation can be calculated using the receptor affinities (see Methods, Eq. 1). We report the resulting receptor occupation as the concentration of D1Rs and D2Rs bound to DA (denoted as $[D1-DA]$ and $[D2-DA]$, respectively). Expressing receptor occupation in terms of concentration (typically in nM) follows from our estimates based on experimental measurements (Richfield et al., 1989, 1987), and is convenient for the calculation of binding rates and equilibria (see Methods).

First, we investigated receptor binding for a range of affinities (**Fig. 1**), reflecting the range of measured values in different experimental studies (Neve and Neve, 1997). Due to the low affinity of D1Rs, at low baseline [DA] only a small fraction of D1 receptors may be occupied. However, there are overall more D1Rs than D2Rs (Richfield et al., 1989), and $\approx 80\%$ of D2Rs are retained in the endoplasmatic reticulum (Prou et al., 2001). Therefore, the concentration of D1Rs in the membrane available to extracellular DA is a lot higher than the concentration of D2Rs (e.g. 20 times more in the nucleus accumbens; Nishikori et al., 1980; see Methods).

Thus, in our simulation, the actual concentration of bound D1Rs ($[D1-DA] \approx 20nM$) was, at DA baseline, much closer to the concentration of bound D2Rs ($[D2-DA] \approx 35nM$) than suggested by the different D1 and D2 affinities alone. We further confirmed that this was not due to a specific choice of the dissociation constants in the model, as $[D1-DA]$ and $[D2-DA]$ remained similar over the range of experimentally measured D1R and D2R affinities (Neve and Neve, 1997) (**Fig. 1a**). This suggests that $[D1-DA]$ is at most twice as high as $[D2-DA]$ instead of 40 times higher as suggested by the difference in fraction of bound receptors. Therefore, $[D1-DA]$ and $[D2-DA]$ might be better indicators for the signal transmitted to MSNs, as the fraction of bound receptors neglects the different receptor type abundances.

Next, we investigated the effect of slow [DA] changes (Grace, 1995; Schultz, 1998; Floresco et al., 2003) by exposing our model to changes in the [DA] baseline. For signalling timescales that are long with respect to the half-life time of the receptors ($t_{slow} >> t_{1/2} \approx 80s$), we used the dissociation constant to calculate the steady state receptor occupancy. We found that for a range of [DA] baselines (mimicking slow changes in [DA]), there was less than a two-fold difference between $[D1-DA]$ and $[D2-DA]$ (**Fig. 1b**), because of the different abundances of D1 and D2 receptors. This is in contrast to affinity-based models, which suggest that D2Rs are better suited to encode slow or tonic changes in [DA]. Interestingly the change of $[D1-DA]$ was almost linear in [DA], while the change of $[D2-DA]$ showed nonlinear effects due to the change in available free D2R. Thus, based on these results, it could even be argued that D1Rs are better at detecting tonic signals at high [DA] levels, since they do not saturate as easily.

While for baseline and slow changes in [DA] the receptor occupation can be determined based on the receptor affinity, fast changes in [DA] also require a description of the underlying receptor kinetics. To investigate the effect of typical DA signals on receptor occupation, we developed a kinetics model incorporating binding and unbinding rates that determine the overall receptor affinity (see Methods, Eq. 8, 9). The available experimental measurements indicate that the different D1R and D2R affinities are largely due to different binding rates, while their unbinding rates are similar (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Richfield et al., 1989). We incorporated these measurements into our kinetics model and investigated the model's response to a variety of fast DA signals.

We started by measuring the model response to a [DA] step change from 20nM to $1\mu M$. This is quite a large change compared to phasic DA signals *in vivo* (Robinson et al., 2001; Cheer et al., 2007; Hamid et al., 2016), which we choose to illustrate that our results are not just due to a small amplitude DA signal. We found that binding to both receptor subtypes increased very slowly. Even for the high affinity D2Rs it took more than $5s$ to reach their new equilibrium (**Fig. 1c**). Thus, unlike the affinity-based model, our model suggests that the D2Rs will not saturate for single reward events, which last overall for up to $\approx 3s$. Note that the non-saturation is independent of the abundance of the receptors and is only determined by the kinetics of the receptors (see Methods). Due to their slow unbinding, D1Rs and D2Rs also took a long time to return to baseline receptor occupancy after a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$ (**Fig. 1d**). Thus, we conclude that with slow kinetics of receptor binding both D1Rs and D2Rs can detect single phasic DA signals and that both remain occupied long after a high [DA] has returned to baseline.

DA receptor binding kinetics for different types of DA signals

Next, we investigated $[D1 - DA]$ and $[D2 - DA]$ for three different types of DA signals (**Fig. 2**). The first signal was a phasic DA increase ('long burst', **Fig. 2**a), mimicking responses to rewards and reward-predicting stimuli (Robinson et al., 2001; Cheer et al., 2007). The second signal was a brief phasic DA increase, followed by a decrease ('burst-pause', **Fig. 2**a), mimicking responses to conditioned stimuli during extinction (Pan et al., 2008) or to other salient stimuli (Schultz, 2016). The third signal was a prolonged DA ramp, mimicking a value signal when approaching a goal (Howe et al., 2013; Hamid et al., 2016) (**Fig. 2**b). In the affinity-based model with instant kinetics the D1Rs mirrored the [DA] time course for all three types of signals, since even at $[DA] = 200nM$ D1Rs are far from saturation. By contrast, D2Rs showed saturation effects as soon as $[DA] > 2 \cdot K_D^{D2}$, leading to differing D1 and D2 time courses (**Fig. 2**, grey traces). Importantly, in our model with slow kinetics, the time courses of $[D1 - DA]$ and $[D2 - DA]$ were nearly identical (**Fig. 2**, bottom row), supporting that both receptor types are equally affected by phasic DA signals. This was the case for all the three signals: burst, burst-pause and ramping DA signals. The only difference between the $[D1 - DA]$ and $[D2 - DA]$ time courses were the absolute amplitudes. For example, $[D2 - DA]$ started from a baseline about twice as high as $[D1 - DA]$, but then also responded

to the long burst DA signal with a change about twice as high. The similarity of $[D1-DA]$ and $[D2-DA]$ responses to both slow (**Fig. 1b**) and fast (**Fig. 2**), [DA] changes indicates that the different DA receptor types respond similarly independent of the timescale of [DA] changes. It could even be argued that D2Rs are better at detecting phasic DA signals, since they respond with a larger absolute change in occupied receptors.

To understand why the D1Rs and D2Rs respond in a similar fashion, we considered the relevant model parameters in more detail. The binding rate constants of D1Rs and D2Rs differ by a factor of ≈ 60 ($k_{on}^{D1} = 0.0003125nM^{-1}min^{-1}$ and $k_{on}^{D2} = 0.02nM^{-1}min^{-1}$; Burt et al., 1976; Sano et al., 1979; Maeno, 1982; see also Methods), suggesting faster D2Rs. However, experimental data indicates that there are ≈ 40 fold more unoccupied D1 receptors ($[D1] \approx 1600nM$) than unoccupied D2 receptors ($[D2] \approx 40nM$) on MSN membranes in the extracellular space of the rat striatum. This difference is due to a combination of simply higher abundance of D1 receptors (Richfield et al., 1987, 1989) and D2 receptors being retained in the endoplasmatic reticulum (Prou et al., 2001) . Therefore, the absolute binding rate $\frac{d[DX-DA]}{dt}^{+} = k_{on} \cdot [DA] \cdot [DX]$ differs only by a factor of ≈ 1.5 between the D1Rs and D2Rs. That is, the difference in the kinetics of D1Rs and D2Rs is compensated by the different receptor numbers, resulting in nearly indistinguishable aggregate kinetics (**Fig. 2**). This is consistent with recent experimental findings that D2R expressing MSNs can detect phasic [DA] signals (Marcott et al., 2014; Yapo et al., 2017).

The dynamics introduced by the slow kinetics in our model also affected the time course of DA signalling. With instant kinetics the maximum receptor occupancy was reached at the peak [DA] (**Fig. 2**). By contrast, for slow kinetics the maximum receptor occupancy was reached when [DA] returned to its baseline (**Fig. 2a**) because as long as [DA] was higher than the equilibrium value of [D1-DA] and [D2-DA], more receptors continued to become occupied. Therefore, for all DA signals, the maximum receptor occupancy was reached towards the end of the pulse.

Another striking effect of incorporating receptor kinetics was that a phasic increase in [DA] kept the receptors occupied for a long time (**Fig. 2**a, green traces). However, when a phasic increase was followed by a decrease, [D1-DA] and [D2-DA] returned to baseline much faster (**Fig. 2**a, purple traces). This indicates that burst-pause firing patterns can be distinguished from pure burst firing patterns on the level of the MSN DA receptor occupancy. Furthermore,

this supports the view that the fast component of the DA firing patterns (Schultz, 2016) is a salience response, and points to the intriguing possibility that the pause following the burst can, at least partly, revoke the receptor-ligand binding induced by the burst. In fact, for each given burst amplitude, a sufficiently long pause duration could cancel the receptor activation (**Fig. 3**), with larger [DA] amplitudes requiring longer pauses to cancel the activation. Thereby, the burst-pause firing pattern of DA neurons could effectively signal a reward "false-alarm".

The long time it took $[D1-DA]$ and $[D2-DA]$ to return to baseline after phase DA signals (**Fig. 2a**) indicates that the receptor occupation integrates DA signals over time. To examine this property, we simulated a sequence of DA signals on a timescale relevant for behavioural experiments (**Fig. 4**). Each sequence consisted of 50 events and each event was separated by 15 s. Three different types of sequences were tested: 50 phasic DA bursts, 40 phasic DA bursts followed by 10 burst-pause signals, and 40 phasic DA bursts followed by 10 non-events. We found that both $[D1-DA]$ and $[D2-DA]$ accumulated over the sequence of DA signals. The sawtooth shape of the curves was due to the initial unbinding of the receptors after each burst event, which was then interrupted by the next DA signal 15 s later. At higher levels of receptor activation, the amount of additional activated receptor per DA pulse was reduced since there are less free receptors available, and the amount of receptors unbinding during the pulse duration was increased because more receptors were occupied. The accumulation occurred as long as the time interval between the DA signals was shorter than $\approx 2 \cdot t_{1/2}$. Together, the shape and period of the DA pulses lead to the formation of an equilibrium, visible here as a plateau for the absolute amount of occupied receptor. This occurred at the level at which the amount of receptors unbinding until the next DA burst was the same as the amount of receptors getting occupied by the DA burst. Finally, the burst-pause events did not lead to an accumulation of occupied receptors over time. In fact, the receptor occupation was the same for burst-pause and non-event, except during the short burst component of the burst-pause events (note the overlapping green and orange curves in **Fig. 4**). This extends the property of burst-pause signals as "false alarm" signals to a wide range of occupancy levels.

Incorporating the slow kinetics in the model is crucial for functional considerations of the DA system. Currently, following the affinity-based model, the amplitude of a DA signal (i.e. peak [DA]) is often considered as a key signal, e.g. in the context of reward magnitude or probability (Morris et al., 2004; Tobler et al., 2005; Hamid et al., 2016). However, as DA unbinds slowly (over tens of seconds; **Fig. 1d**) and the binding rate changes approximately linearly with [DA], the amount of receptor occupancy does not primarily depend on the amplitude of the [DA] signal.

Due to the linearity of the binding rate, the receptor occupation increases linearly with time and $[DA]-[DA]_{baseline}$, while the unbinding is negligible as long as $t << t_{1/2}$. Therefore the integral of the [DA] time course should be a close approximation of the receptor occupation for signals that are shorter than the half-life time of the receptors. We confirmed this consideration by simulating a range of DA signals (burst, burst-pauses, and ramps) with different durations and amplitudes. For each DA signal we compared its area under the curve with the resulting peak change in the absolute receptor occupancy. For both D1R and D2R we found that the maximum receptor activation was proportional to the area under the curve of the [DA] signal, while independent of its specific time course (**Fig. 5**). The small deviation from the proportionality seen for large-area DA signals for the D2Rs was due to the decrease in the amount of free receptor as more and more receptors were bound. In this regime the assumption that the binding rate is linear with [DA] was slightly violated leading to the non-proportionality. In contrast, the relationship between the DA burst amplitude and resulting receptor occupation was fit well by a quadratic function (**Fig. 5b**), which reflects quadratic increases in the area under the curve for larger amplitudes.

The overall striking proportionality of the integral of the DA signal with receptor binding indicates that D1Rs and D2Rs act as slow integrators of the DA signal. Interestingly, this means that DA ramps, even with a relatively small amplitude (**Fig. 2b**), are an effective signal to occupy DA receptors. In contrast, for locally very high [DA] (e.g. at corticostriatal synapses during phasic DA cell activity; Grace et al., 2007) our model predicts that the high concentration gradient would only lead to a very short duration of this local DA peak and thereby make it less effective in occupying DA receptors.

To further test the generality of our findings, we examined our model responses systematically for a set of different DA time courses (**Fig. 6** and **Fig. 7**). While the shape of the DA pulses strongly affected the time courses of the receptor activation, D1 and D2 receptor activation were virtually identical for a given pulse shape. For DA bursts with different amplitudes (**Fig. 6a**), higher amplitudes of the DA burst lead to stronger receptor activation. However, the relationship between burst amplitude and receptor occupation was not linear, but instead reflected the area under the curve of the DA pulse (see above).

Importantly, despite 'slow' kinetics, the onset of the increase in $[D1-DA]$ and $[D2-DA]$ is immediate and determined by the area under the curve of the DA pulse up to each time point. This means that the DA receptor occupation reflects an ongoing integration of the DA signal. Furthermore, this provides us with an intuition for how the DA receptor occupation develops during a particular DA signal. The increase of the receptor occupation is controlled by the DA pulse shape and is proportional to the area under the curve of the DA pulse up to each time point. So with realistic kinetics the receptor occupation will always reach its maximum towards the end of the DA pulse and its rise profile depends on the specific pulse shape.

For a fixed burst amplitude, we also determined the effect of different DA re-uptake rates to look at potential differences in DA signalling in dorsal and ventral striatum, with fast and slow re-uptake, respectively. This was done by changing the parameter V_{max} (see Methods), which controlled the time the [DA] took to return to the baseline from the peak value (**Fig. 6b**). While this had only a small visible effect on the input DA signal (**Fig. 6b**, top panel), the resulting $[D1-DA]$ and $[D2-DA]$ were quite different. The difference in the resulting occupancy levels can again be understood in terms of the role of the pulse shape discussed above. For fast uptake, i.e. high V_{max}, DA is cleared faster from the system which reduces the area under the curve of the DA pulse but does not affect the peak DA levels during the burst. This leads to an important distinction between the affinity-based model and the model with realistic kinetics. With realistic kinetics the resulting receptor occupancy is lower for higher V_{max} because the reduced area under the DA pulse gives the receptors less time to get occupied during high [DA] conditions (**Eq. 5**). By contrast, this property is not seen in the affinity-based model, in which the time course of $[D1-DA]$ and $[D2-DA]$ simply follows the input [DA] signal, and thereby, the peak receptor occupation levels are not affected by

480 V_{max}.

481 Next, we examined DA ramps with different time courses, but the same maximal amplitude.
482 Again, consistent with our consideration of the important role of the area under the curve
483 of DA signals, we found that longer ramps lead to larger DA receptor occupation (**Fig. 7a**).
484 We then investigated the DA signals that included the effects of pauses in DA cell activity
485 further. First, we tested burst-pause signals and determined the role of the duration of the
486 pause, when the [DA] decayed to zero. For a fixed burst amplitude and duration, a different
487 duration of the subsequent pause lead to differing receptor activation levels when the burst-
488 pause signal was over (**Fig. 7b**). This indicates that DA pauses are very effective in driving
489 the receptor occupation quickly back to baseline (i.e. within few seconds) because, in this
490 case, the receptor occupation changes reflect solely the unbinding rates. In contrast, for a
491 burst followed by a return to baseline [DA], the decrease in receptor occupation would be
492 slower because during the baseline portion of the signal both binding and unbinding processes
493 play a role, and the binding counteracts some of the unbinding (see **Eq. 5** and **Eq. 6**).

494 In this context we also looked at pure DA pauses (i.e. without a preceding burst), e.g. reflecting
495 DA cell responses to aversive stimuli (Schultz, 2007) that lead to reductions in [DA] (Roitman
496 et al., 2008). These signals also lead to fast decreases in [D1-DA] and [D2-DA], with the
497 duration of the pause having a strong effect on the amplitude and duration of the decrease
498 (**Fig. 7c**). Pauses in DA cell firing may not necessarily lead to a [DA] of zero, as e.g. local
499 mechanisms of DA release at terminals may persist even when DA neurons do not spike.
500 Therefore, we also examined a scenario in which the DA pauses involved a reduction of [DA]
501 to a quarter of its baseline level (Owesson-White et al., 2012). We found that in this case too,
502 the pauses were effective in driving the receptor occupation back to the baseline (**Fig. 7c**,
503 dashed line). However, the net unbinding was slower than during conditions when $[DA]$
504 was zero during the pause. This is consistent with our corresponding observations for the
505 burst-pause signals above (**Fig. 3**).

506 D1R and D2R occupancy in a probabilistic reward task paradigm

507 A general effect of the slow kinetics was that DA receptors remained occupied long after the
508 DA pulse was over (**Fig. 2**), so that the effect of DA pulses was integrated over time (**Fig. 4**).

To investigate the information that is preserved in the receptor occupation about DA signals on time scales relevant for behavioural tasks, we simulated sequences with probabilistic DA events (see Methods). First, we compared sequences, in which every $15 \pm 5s$ there was a DA burst with either 30%, 50%, or 70% probability (**Fig. 8a, c**). The resulting changes in $[D1 - DA]$ and $[D2 - DA]$ confirmed the integration of DA pulses over minutes. The integration of DA bursts was due to DA bursts arriving before the receptor occupation caused by the previous pulses had decayed, leading to an increased receptor activation compared to single DA bursts (**Fig. 4**). We then examined whether the DA receptor occupancy can distinguish different reward probabilities by using a simple classifier comparing two sequences with each other (see Methods). We tested sequences from 0% to 100% probability in steps of 10%, and ordered the resulting classification success in terms of the difference in reward probability between the two sequences (**Fig. 8b, d, e, f**). For example, a comparison between a 30% and a 70% reward probability sequence yields a data point for a 40% difference. For both D1 and D2 receptors, we found that already for differences of 10% the classification exceeded chance level, and yielded near perfect classification around a 40% difference. Overall, the classification was slightly better for D1R due to their slightly less developed plateauing response (**Fig. 4**). The successful classification of reward probabilities demonstrates that it would be possible for striatal neurons to read out different reward rates from DA receptor occupancy in a behavioural task. The classification in this example is only possible because of the slow kinetics of the receptors and, importantly, not due to an accumulation of DA. In this simulation there was no accumulation of DA itself and the [DA] dropped back to baseline in between the DA events. This provides a potential neural substrate for how fast DA signals can lead to an encoding of the slower reward rate, which can be utilized as a motivational signal (Mohebi et al., 2019).

Validation for fast binding kinetics

Our model assumption of slow kinetics was based on neurochemical estimates of wildtype DA receptors (Burt et al., 1976; Sano et al., 1979; Maeno, 1982). In contrast, recently developed genetically-modified DA receptors, used to probe [DA] changes, have fast kinetics (Sun et al., 2018; Patriarchi et al., 2018). Although the kinetics of the genetically modified DA receptors are unlikely to reflect the kinetics of the wildtype receptors (see Discussion), we also examined the effect of faster DA kinetics in our model. Fast kinetics were implemented by multiplying

k_{on} and k_{off} by a factor q, keeping K_D constant. We found that the similarity between $[D1-DA]$ and $[D2-DA]$ persists even if the actual kinetics were a 100 times faster than assumed in our model (**Fig. 9**). This was the case for all types of [DA] signals because the difference between the aggregate D1 and D2 binding rates (Eq. 5) still only differed by a factor of 1.5. Furthermore, the D2Rs did not show visible saturation effects even for $q = 100$. Faster kinetics mostly affected the amplitude of the receptor response and the time it took to return to baseline receptor occupancy. However, only for $q = 100$ the receptor occupation dropped slightly below baseline during the pauses of a burst-pause DA signal (**Fig. 9a, b**). On a longer time scale with repetitive DA bursts (**Fig. 9e, f**) D1Rs and D2Rs integrated the DA bursts over time even if kinetics were twice as fast ($q = 2$). This is because the half-time of the receptors were 40 s (for $q = 2$), while the DA burst signal was repeated every 15 s. Thereby, [D1-DA] and [D2-DA] were dominated by the repetition of the signal rather than by the impact of individual DA burst signals. In contrast, for $q = 10$ the change in receptor occupancy was dominated by the single pulses, since the half-life time was 8s, whereby the receptors mostly unbind in between DA pulses. Therefore, our results concerning the similarity of D1 and D2 receptors do not depend on the exact kinetics parameters or potential temperature effects, as long as the parameter changes are roughly similar for D1 and D2 receptors. However, DA receptor kinetics faster by a factor of 10 or more affected the ability of DA receptor occupancy to integrate DA pulses over time (**Fig. 9e, f**).

In our model we assumed homogeneous receptor populations, namely that all D1 receptors have a low affinity and that all D2 receptors have a high affinity. However, this could be a simplification, as $\approx 10\%$ of D2 receptors may exist in a low affinity state, while $\approx 10\%$ of the D1 receptors may be in a high affinity state (Richfield et al., 1989). Therefore, we also incorporated different affinity states of the D1 and D2 receptors into our model. The D1Rs in a high affinity state ($D1^{high}$) were modelled by increasing the on-rate of the D1R but keeping its off-rate constant, creating a receptor identical to the $D2^{high}$ receptor. Although the high affinity state kinetics of the D1R are currently unknown, we choose this model as a faster on-rate potentially has the strongest effect on our conclusions. Correspondingly, we modelled the $D2^{low}$ receptor as a D2R with slower on-rate, which was equivalent to simply reducing $[D2^{tot}]$ since the $D2^{low}$ receptors were predominantly unoccupied during baseline DA and bound only sluggishly to DA during phasic signals. The main effect of incorporating the different receptor affinity states was a change in the respective equilibrium values of absolute concentration of

receptors bound to DA (**Fig. 10**). However, importantly, taking into account these different affinity states, preserved the similarity of time courses of D1R and D2R occupancy and the ability to integrate DA pulses over time (**Fig. 10** and **Fig. 11**) since the half-life time of both receptors remained long.

Discussion

The functional roles of DA in reward-related learning and motivation have typically been studied by characterizing the firing patterns of dopaminergic neurons and the resulting changes in striatal [DA] (Schultz, 2007). In contrast to other, more conventional neurotransmitters like glutamate or GABA, the release of DA in the striatum may form a global signal that affects large parts of the striatum similarly (Schultz, 1998). Such global [DA] changes involve longer time scales lasting at least several seconds (Roitman et al., 2008; Howe et al., 2013). Importantly, to affect neural activity in the striatum, DA first needs to bind to DA receptors. This process is often simplified by assuming that this happens instantaneously, so that every change in [DA] is immediately translated into a change in DA receptor occupation. As this contradicts physiological measurements of the receptor kinetics (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Nishikori et al., 1980), we developed and investigated a model incorporating DA receptor kinetics as well as differences in D1 and D2 receptor abundance in the striatum.

Our results cast doubt on several long-held views on DA signalling. A common view is that D1 and D2 MSNs in the striatum respond to different DA signals due to the affinity of their predominant receptor type. Accordingly, phasic DA changes should primarily affect D1 MSNs, while slower changes or DA pauses should primarily affect D2 MSNs. In contrast, our model indicates that both D1R and D2R systems can detect [DA] changes, independent of their timescale, equally well. That is, slow tonic changes in [DA], phasic responses to rewards, and ramping increases in [DA] over several seconds lead to a similar time course in the response of D1 and D2 receptor occupation in our model. However, the baseline level of activated DA receptors and the amplitude of the response was typically twice as high in D2 compared to D1 receptors. Although, D1 and D2 receptors have opposing effects on the excitability (Flores-Barrera et al., 2011) and strength of cortico-striatal synapses (Centonze et al., 2001),

we challenge the view that differences in receptor affinity introduce additional asymmetries in D1 and D2 signalling. Instead of listening to different components of the DA signal, D1 and D2 MSNs may respond to the same DA input. This would actually increase the differential effect on firing rate responses of D1 and D2 MSNs because the opposite intracellular effects of D1 and D2 activation (Surmeier et al., 2007) occur then for the whole range of DA signals.

Recently, ramps in [DA], increasing over several seconds towards a goal, have been connected to a functional role of DA in motivation (Howe et al., 2013; Hamid et al., 2016). In our model DA ramps were very effective in occupying DA receptors due to their long duration. In contrast, for brief phasic increases, the receptor occupation was less pronounced. Overall, our model predicts that the area under the curve of DA signals determines the receptor activation, which puts more emphasis on the duration of the signals, rather than the amplitude of brief DA pulses.

Our model is also relevant for the interpretation of burst-pause firing patterns in DA neurons. These are a different firing pattern than the typical reward-related bursts, and consist of a brief burst followed by a brief pause in action potentials. Such two-component responses of DA cells may reflect saliency and value components, respectively (Schultz, 2016). For example, during extinction learning burst-pause firing patterns have been observed as a response to conditioned stimuli, with each component lasting about 100 ms (Pan et al., 2008). Our model provides a mechanistic account for how the burst-pause DA signals have a different effect on MSNs than pure burst signals, which is important to distinguish potential rewarding signals from other salient, or even aversive stimuli. In our model the pause following the burst was very effective in reducing the number of occupied receptors quickly, thereby preventing the otherwise long-lasting receptor occupation due to the burst. Thereby, canceling the effect of the brief burst might be a neural mechanism to correct a premature burst response that was entirely based on saliency rather than stimulus value (Schultz, 2016). As fast responses of DA cells to potentially rewarding stimuli are advantageous to quickly redirect behaviour, the subsequent pause signal might constitute an effective correction mechanism labelling the burst as a false alarm.

Functionally, the slow unbinding rate of D1 and D2 receptors pointed to an interesting property in integrating phasic DA events over time. The unbinding rate might be one of the mechanisms translating fast DA signals into a slower time scale, which could be a key mechanism to

generate motivational signals (Mohebi et al., 2019). Importantly, the slow kinetics of receptor binding do not prevent a fast neuronal response to DA signals. In our model [DA] changes affected the number of occupied receptors immediately; it just took seconds or even minutes until the new equilibrium was reached. However, reaching the new equilibrium is not necessarily relevant on a behavioural level. Instead the intracellular mechanisms that react to the receptor activation need to be considered to determine which amount of receptor activation is required to affect neural activity. In our model changes on a nanomolar scale occurred within 100 ms, a similar timescale as behavioural effects of optogenetic DA manipulations (Hamid et al., 2016).

The slower time scales were introduced into our model by the kinetics based on in-vitro measurements (Burt et al., 1976; Sano et al., 1979; Maeno, 1982; Nishikori et al., 1980). A limitation of our model is the uncertainty about the accuracy of these measurements, and whether they reflect in-vivo conditions. We addressed this here by also examining faster kinetics, for which there is currently no direct evidence in the literature. However, recently DA receptors have been genetically modified to serve as sensors for fast [DA] changes (Patriarchi et al., 2018), which suggests possible fast kinetics. It seems unlikely though that the kinetics of the genetically-modified receptors represent the kinetics of the wildtype DA receptors, as e.g. the screening procedure to find suitable receptor variants yielded a large range of different affinities (meaning changes in the kinetics of binding, unbinding or both) based on changes at the IL-3 site (Patriarchi et al., 2018). Changes in the IL-3 site have also previously been shown to strongly affect the receptor affinity (Robinson et al., 1994).

In addition to the receptor kinetics, the different abundances of D1 and D2 receptors are also key parameters in our model, which we estimated based on previous experimental studies. However, in case our estimates of the receptor abundances were incorrect, the receptor occupation would still be determined by the kinetic parameters and it would differ substantially from instant kinetics assumed in the affinity-based model. In particular, both receptor types would still not saturate during DA pulses, but integrate the DA signal over longer time scales. Furthermore, our results do not depend on the exact absolute receptor abundances, but on the relative abundances of D1 and D2 receptors. Therefore, our results on the similarity of the D1 and D2 responses hold as long as the abundance of the D1 receptors is roughly an order of magnitude higher than the abundance of the D2 receptors. Overall, we conclude that

it is important to consider the effect of the receptor kinetics on DA signalling, which have not received much attention in experimental studies, nor in theoretical considerations of DA function thus far.

References

Atcherley CW, Wood KM, Parent KL, Hashemi P, Heien ML (2015) The coaction of tonic and phasic dopamine dynamics. *Chemical Communications* 51:2235–2238.

Banay-Schwartz M, Kenessey A, DeGuzman T, Lajtha A, Palkovits M (1992) Protein content of various regions of rat brain and adult and aging human brain. *Age* 15:51–54.

Bergstrom BP, Garris PA (2003) 'passive stabilization'of striatal extracellular dopamine across the lesion spectrum encompassing the presymptomatic phase of parkinson's disease: a voltammetric study in the 6-ohda-lesioned rat. *Journal of neurochemistry* 87:1224–1236.

Berke JD (2018) What does dopamine mean? *Nature neuroscience* p. 1.

Borland LM, Shi G, Yang H, Michael AC (2005) Voltammetric study of extracellular dopamine near microdialysis probes acutely implanted in the striatum of the anesthetized rat. *Journal of neuroscience methods* 146:149–158.

Burt DR, Creese I, Snyder SH (1976) Properties of [3h] haloperidol and [3h] dopamine binding associated with dopamine receptors in calf brain membranes. *Molecular pharmacology* 12:800–812.

Centonze D, Picconi B, Gubellini P, Bernardi G, Calabresi P (2001) Dopaminergic control of synaptic plasticity in the dorsal striatum. *European journal of neuroscience* 13:1071–1077.

Cheer JF, Aragona BJ, Heien ML, Seipel AT, Carelli RM, Wightman RM (2007) Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. *Neuron* 54:237–244.

Copeland RA (2004) *Enzymes: a practical introduction to structure, mechanism, and data analysis* John Wiley & Sons.

688 Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dy-
689 namic shifts in dopamine signaling in the nucleus accumbens. *Nature neuroscience* 10:1020.

690 Day M, Wokosin D, Plotkin JL, Tian X, Surmeier DJ (2008) Differential excitability and mod-
691 ulation of striatal medium spiny neuron dendrites. *Journal of Neuroscience* 28:11603–11614.

692 DiResta G, Lee J, Lau N, Ali F, Galicich J, Arbit E (1990) Measurement of brain tissue density
693 using pycnometry In *Brain Edema VIII*, pp. 34–36. Springer.

694 Dreyer JK (2014) Three mechanisms by which striatal denervation causes breakdown of
695 dopamine signaling. *Journal of Neuroscience* 34:12444–12456.

696 Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD (2010) Influence of phasic and tonic
697 dopamine release on receptor activation. *Journal of Neuroscience* 30:14273–14283.

698 Dreyer JK, Hounsgaard J (2013) Mathematical model of dopamine autoreceptors and uptake
699 inhibitors and their influence on tonic and phasic dopamine signaling. *Journal of neuro-*
700 *physiology* 109:171–182.

701 Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from
702 actions to habits to compulsion. *Nature neuroscience* 8:1481.

703 Fiorillo CD, Newsome WT, Schultz W (2008) The temporal precision of reward prediction in
704 dopamine neurons. *Nature neuroscience* 11:966.

705 Flores-Barrera E, Vizcarra-Chacón BJ, Bargas J, Tapia D, Galarraga E (2011) Dopaminergic
706 modulation of corticostriatal responses in medium spiny projection neurons from direct and
707 indirect pathways. *Frontiers in systems neuroscience* 5:15.

708 Floresco SB, West AR, Ash B, Moore H, Grace AA (2003) Afferent modulation of dopamine
709 neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature neu-*
710 *roscience* 6:968.

711 Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human
712 cognition: psychopharmacological studies with cabergoline and haloperidol. *Behavioral*
713 *neuroscience* 120:497.

Grace AA (1995) The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. *Drug & Alcohol Dependence* 37:111–129.

Grace AA, Floresco SB, Goto Y, Lodge DJ (2007) Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in neurosciences* 30:220–227.

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2016) Mesolimbic dopamine signals the value of work. *Nature neuroscience* 19:117.

Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500:575–579.

Justice Jr J (1993) Quantitative microdialysis of neurotransmitters. *Journal of neuroscience methods* 48:263–276.

Maeno H (1982) Dopamine receptors in canine caudate nucleus. *Molecular and cellular biochemistry* 43:65–80.

Marcott PF, Mamaligas AA, Ford CP (2014) Phasic dopamine release drives rapid activation of striatal d2-receptors. *Neuron* 84:164–176.

Mohebi A, Pettibone JR, Hamid AA, Wong J, Vinson LT, Patriarchi T, Tian L, Kennedy RT, Berke JD (2019) Dissociable dopamine dynamics for learning and motivation. *Nature* .

Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 43:133–143.

Neve KA, Neve RL (1997) Molecular biology of dopamine receptors In *The dopamine receptors*, pp. 27–76. Springer.

Nishikori K, Noshiro O, Sano K, Maeno H (1980) Characterization, solubilization, and separation of two distinct dopamine receptors in canine caudate nucleus. *Journal of Biological Chemistry* 255:10909–10915.

Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191:507–520.

741 Owesson-White CA, Roitman MF, Sombers LA, Belle AM, Keithley RB, Peele JL, Carelli RM,
742 Wightman RM (2012) Sources contributing to the average extracellular concentration of
743 dopamine in the nucleus accumbens. *Journal of neurochemistry* 121:252–262.

744 Pan WX, Schmidt R, Wickens JR, Hyland BI (2005) Dopamine cells respond to predicted
745 events during classical conditioning: evidence for eligibility traces in the reward-learning
746 network. *Journal of Neuroscience* 25:6235–6242.

747 Pan WX, Schmidt R, Wickens JR, Hyland BI (2008) Tripartite mechanism of extinction
748 suggested by dopamine neuron activity and temporal difference model. *Journal of Neuro-*
749 *science* 28:9619–9631.

750 Patriarchi T, Cho JR, Merten K, Howe MW, Marley A, Xiong WH, Folk RW, Broussard GJ,
751 Liang R, Jang MJ et al. (2018) Ultrafast neuronal imaging of dopamine dynamics with
752 designed genetically encoded sensors. *Science* p. eaat4422.

753 Prou D, Gu WJ, Le Crom S, Vincent JD, Salamero J, Vernier P (2001) Intracellular retention of
754 the two isoforms of the d 2 dopamine receptor promotes endoplasmic reticulum disruption.
755 *Journal of Cell Science* 114:3517–3527.

756 Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y,
757 DeLong MR, Obeso JA (2010) Goal-directed and habitual control in the basal ganglia:
758 implications for parkinson's disease. *Nature Reviews Neuroscience* 11:760–772.

759 Reyes BA, Pendergast JS, Yamazaki S (2008) Mammalian peripheral circadian oscillators are
760 temperature compensated. *Journal of biological rhythms* 23:95–98.

761 Reynolds JN, Hyland BI, Wickens JR (2001) A cellular mechanism of reward-related learning.
762 *Nature* 413:67.

763 Richfield EK, Penney JB, Young AB (1989) Anatomical and affinity state comparisons between
764 dopamine d 1 and d 2 receptors in the rat central nervous system. *Neuroscience* 30:767–777.

765 Richfield EK, Young AB, Penney JB (1987) Comparative distribution of dopamine d-1 and
766 d-2 receptors in the basal ganglia of turtles, pigeons, rats, cats, and monkeys. *Journal of*
767 *Comparative Neurology* 262:446–463.

Richfield EK, Young AB, Penney JB (1989) Comparative distributions of dopamine d-1 and d-2 receptors in the cerebral cortex of rats, cats, and monkeys. *Journal of Comparative Neurology* 286:409–426.

Robinson DL, Phillips PE, Budygin EA, Trafton BJ, Garris PA, Wightman RM (2001) Sub-second changes in accumbal dopamine during sexual behavior in male rats. *Neuroreport* 12:2549–2552.

Robinson SW, Jarvie KR, Caron MG (1994) High affinity agonist binding to the dopamine d3 receptor: chimeric receptors delineate a role for intracellular domains. *Molecular pharmacology* 46:352–356.

Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM (2004) Dopamine operates as a subsecond modulator of food seeking. *Journal of Neuroscience* 24:1265–1271.

Roitman MF, Wheeler RA, Wightman RM, Carelli RM (2008) Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nature neuroscience* 11:1376.

Sano K, Noshiro O, Katsuda K, Nishikori K, Maeno H (1979) Dopamine receptors and dopamine-sensitive adenylate cyclase in canine caudate nucleus: Characterization and solubilization. *Biochemical pharmacology* 28:3617–3627.

Schultz W (1998) Predictive reward signal of dopamine neurons. *Journal of neurophysiology* 80:1–27.

Schultz W (2007) Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* 30:259–288.

Schultz W (2016) Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience* 17:183.

Suaud-Chagny M, Chergui K, Chouvet G, Gonon F (1992) Relationship between dopamine release in the rat nucleus accumbens and the discharge activity of dopaminergic neurons during local in vivo application of amino acids in the ventral tegmental area. *Neuroscience* 49:63–72.

795 Sun F, Zeng J, Jing M, Zhou J, Feng J, Owen SF, Luo Y, Li F, Wang H, Yamaguchi T et al.
796 (2018) A genetically encoded fluorescent sensor enables rapid and specific detection of
797 dopamine in flies, fish, and mice. *Cell* 174:481–496.

798 Surmeier DJ, Ding J, Day M, Wang Z, Shen W (2007) D1 and d2 dopamine-receptor mod-
799 ulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in*
800 *neurosciences* 30:228–235.

801 Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME (2016) Action initiation
802 shapes mesolimbic dopamine encoding of future rewards. *Nature neuroscience* 19:34.

803 Syková E, Nicholson C (2008) Diffusion in brain extracellular space. *Physiological re-*
804 *views* 88:1277–1340.

805 Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine
806 neurons. *Science* 307:1642–1645.

807 Venton BJ, Zhang H, Garris PA, Phillips PE, Sulzer D, Wightman RM (2003) Real-time
808 decoding of dopamine concentration changes in the caudate–putamen during tonic and
809 phasic firing. *Journal of neurochemistry* 87:1284–1295.

810 Yapo C, Nair AG, Clement L, Castro LR, Hellgren Kotaleski J, Vincent P (2017) Detec-
811 tion of phasic dopamine by d1 and d2 striatal medium spiny neurons. *The Journal of*
812 *physiology* 595:7451–7475.

Figure 1: Baseline levels of D1 and D2 receptor occupation and impact of slow kinetics. **(a)** Equilibrium values of absolute concentration of receptors bound to DA as a function of receptor affinities. Here, baseline [DA] was fixed at 20 nM. **(b)** Equilibrium values of absolute concentration of receptors bound to DA as a function of baseline [DA]. Here $K_D^{D1} = 1.6\mu M$ and $K_D^{D2} = 25nM$. '$\times$' and '$+$' indicate the model default parameters. Coloured bands mark the range of values for up to $\pm 20\%$ different receptor abundances. **(c)** Temporal dynamics of D1 and D2 receptor occupancy for a large step up from $[DA] = 20nM$ to $[DA] = 1\mu M$ at time $t = 0$. The gray dotted line shows the D2 equilibrium value (EQM). **(d)** Same as in **c** but for a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$.

Figure 2: Impact of receptor kinetics on responses to different DA signals. **(a)** Two different DA signals, long burst and burst-pause, were simulated. The top panel shows the time course of the model [DA] input signal, and the resulting changes in D1 and D2 receptor occupancy are shown in the other panels below. In the two middle panels we compare [DA-D1] and [DA-D2] in the realistic kinetics model (colored traces, left scales) with the affinity-based model (dashed gray traces, right scales). The time points with maximum receptor occupancy (marked with '$\times$' and 'o' for D1 and D2, respectively) coincided for instant kinetics (purple symbols) with the [DA] peak (combined x and o in top panel), while for slow kinetics (black symbols) it coincided with the offset of the [DA] signal instead (combined x and o in top panel). In the bottom panel [DA-D1] and [DA-D2] are normalized with respect to their baseline at time zero (NU, normalized units). The relative changes with respect to the baseline levels were nearly identical for D1 and D2 receptors. **(b)** Same as in (a) but for ramping DA signals.

Figure 3: Burst-pause DA signals (top panel) did not lead to a prolonged D1 or D2 receptor occupation (middle and bottom panels, respectively). The initial increase in receptor occupation due to the burst component was quickly cancelled by the unbinding that occurred during the pause component. Higher burst amplitudes required a longer pause duration for the cancellation of the receptor occupation. This effect occurred also when the [DA] during the pause did not decay to zero, but to a quarter of the baseline (QBL) [DA] instead. In this case the pause duration had to be longer than for [DA] decaying to zero in order to reset the receptor occupation.

Figure 4: D1 and D2 receptor occupation integrates DA signals over a behavioural time scale. **(a)** The absolute receptor occupancy for D1Rs for three different types of sequences consisting of 50 DA events each. The sequences consisted of 50 long burst events (blue), 40 long burst followed by 10 burst-pause events (orange) and 40 long burst events followed by 10 non-events (green, for comparison). **(b)** Same as in **a** but for D2Rs. Note that for the time course of the overall receptor occupation burst-pause signals are basically identical to non-events.

Figure 5: DA receptor occupation is proportional to the area under the curve of DA signals. **(a)** The peak change in absolute receptor occupancy of D1Rs and D2Rs increases linearly with the area under the curve of the DA pulses. Each data point provides the result of a single simulation in which the indicated parameter (burst amplitude $\Delta[DA]^{max}$, ramp rise time t_{rise}, and DA re-uptake rate V_{max}) was varied (see Methods for parameter values). **(b)** The peak change in absolute receptor occupancy increases non-linearly with the peak [DA] of a phasic burst ($\Delta[DA]^{max}$) (shown here for D1Rs). Data points from single simulations with different $\Delta[DA]^{max}$ (marked by '$\times$') are fit well by a quadratic, but not a linear function. The quadratic fit matches the values obtained from the area under the curve of the DA signal from (a).

Figure 6: Parameter exploration for phasic DA bursts (top row) with the resulting changes in D1 (middle row) and D2 (bottom row) receptor occupancy. **(a)** Effect of variations in the amplitude $\Delta[DA]^{max}$ of the phasic DA burst (top row) on the D1 (middle row) and D2 (bottom row) receptor occupancy. **(b)** Effect of change in the re-uptake rate V_{max} rate (top row) on the D1 (middle row) and D2 (bottom row) receptor occupancy. V_{max} was changed to mimic conditions for the ventral and dorsal striatum. Blue circles and black crosses mark the time points of maximum receptor occupancy for D1 and D2, respectively. Note that for both D1R and D2R the time of maximum receptor occupancy was near the end of the DA signal and that D1Rs and D2Rs behaved similarly independent of the specific parameters of the DA pulse.

Figure 7: Parameter exploration for different DA signals (top row) with the resulting changes in D1 (middle row) and D2 (bottom row) receptor occupancy. **(a)** D1 (middle row) and D2 (bottom row) receptor occupancy for different rise time t_{rise} of the DA ramps (top row). The rise time controls the amount and duration of D1 (middle row) and D2 (bottom row) receptor occupancy. **(b)** D1 (middle row) and D2 (bottom row) receptor occupancy for different pause duration t_{pause} of the burst-pause type DA signals (top row). **(c)** D1 (middle row) and D2 (bottom row) receptor occupancy for different pause duration t_{pause} of DA pauses (without a preceding burst). Such a DA pause led to a fast reduction of receptor occupancy, which took 10s of seconds to return to baseline. The inset shows an enlarged version of the [DA] time course. In one simulation the [DA] is not set to zero during the pause, but to a quarter of the baseline [DA] instead (QBL). The blue circles and black crosses mark the time points of maximum receptor occupancy for D1 and D2, respectively (a-b), or of minimal receptor activation (c). Note that for both D1R and D2R the time of maximum (or minimum for c) receptor occupancy was near the end of the DA signal and that D1Rs and D2Rs behaved similarly independent of the specific parameters of the DA pulse.

Figure 8: Encoding of reward rate by integration of DA signals over minutes in a simulation of a behavioural task. **(a)** Time course of D1 receptor occupancy for sequences of 50 trials with a reward probability, as indicated, in each trial. **(b)** True and false positive rates of the difference in reward probability based on the D1 and D2 receptor occupancy by a simple classifier. Each dot indicates the true and false positive rate from a simulation scenario with the difference in reward probability indicated by the colour. The colour indicates the difference in reward probability (e.g. a 10% difference in purple occurs for 80% vs. 90%, 70% vs. 80%, etc.), and the squares denote the corresponding averages. The red line indicates chance level performance, and a perfect classifier would be at 1.0 true and 0.0 false positive rate. **(c, d)** The same as in panels **a** and **b** but for D2 receptors. **(e)** True positive rates for the classification in a sample session (70% vs 30% reward probability) based on the receptor occupancy of D1 (orange) and D2 (blue) receptors. After a short "swing-in" the receptors distinguished between a 70% and a 30% reward rate. **(f)** Accuracy of the classifier for a range of reward probability differences for the D1 (orange) and D2 (blue) receptors for individual sessions and corresponding session averages.

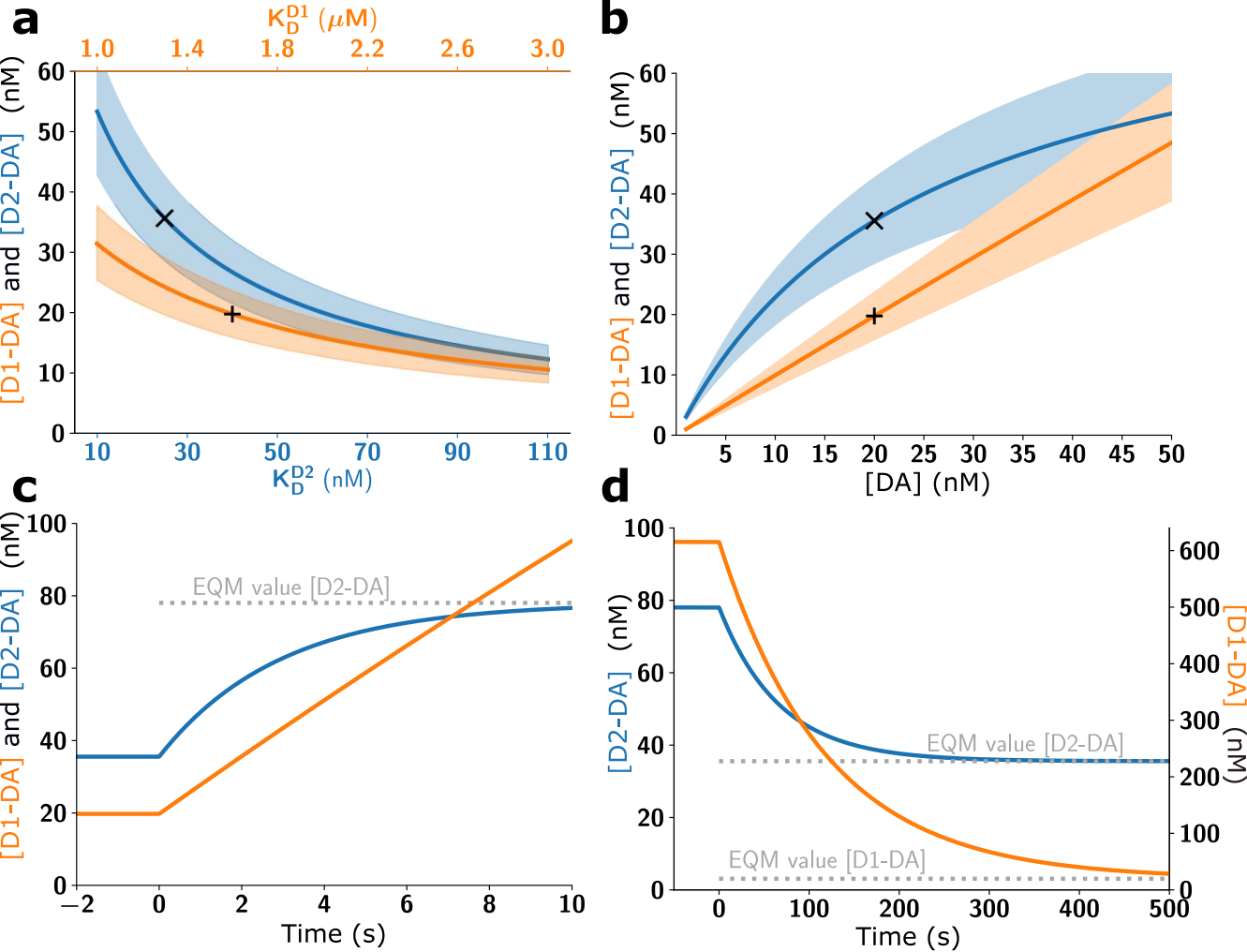
Figure 9: Similarities between D1 and D2 responses persist even if kinetics are much faster than our estimate. Absolute D1R occupancy ([D1-DA]; left column) and D2R occupancy ([D2-DA]; right column) were examined for burst-pause DA signals **(a, b)**, burst-only DA signals **(c, d)**, and the behavioural sequence **(e, f)** (i.e. same simulation scenarios as in Fig. 2a and the 50 bursts pattern from Fig. 4).

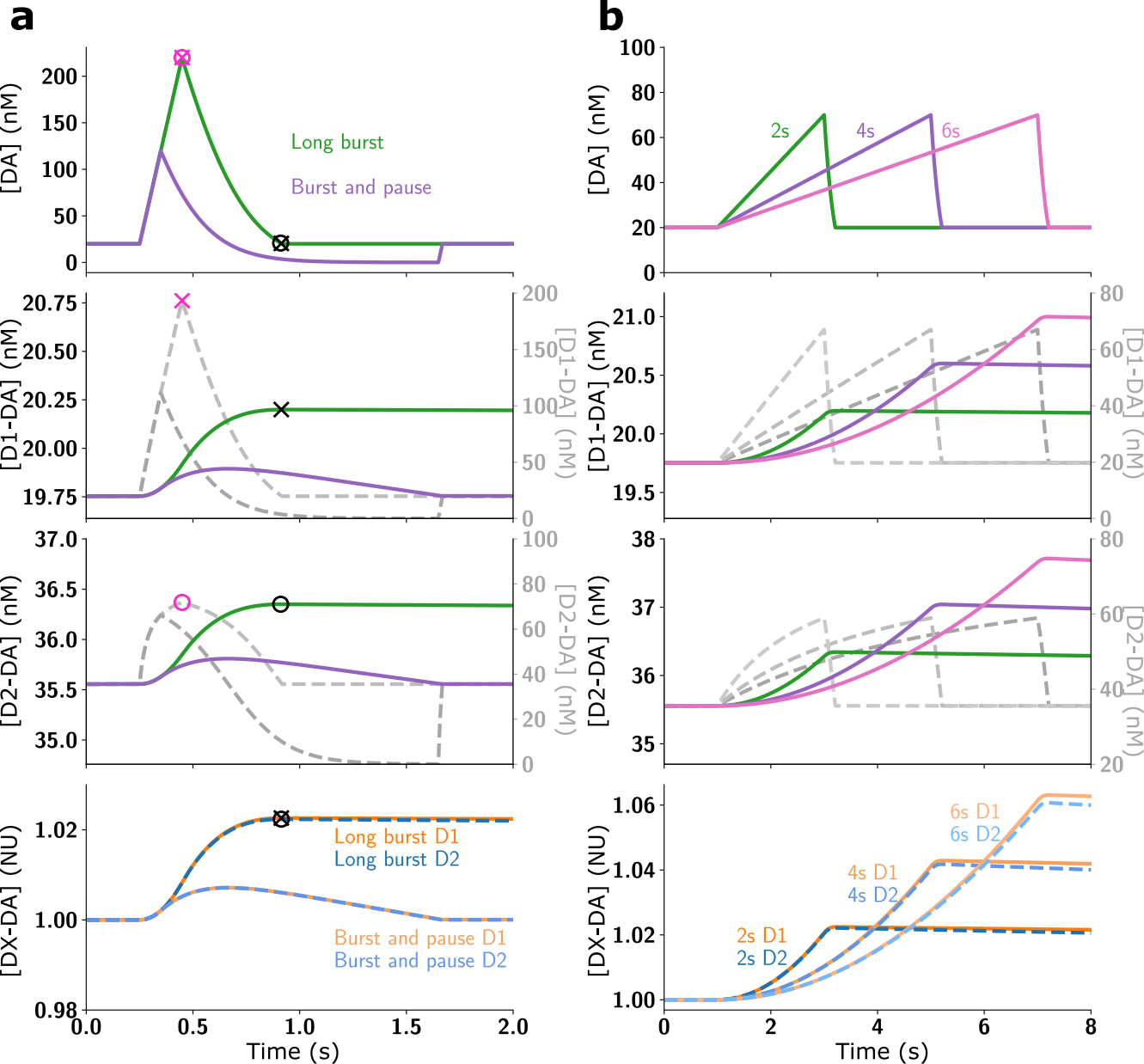
Figure 10: Baseline levels of D1 and D2 receptor occupation and impact of slow kinetics with different receptor affinity states. Here 10% of D1R are assumed to be in a high affinity state ($D1^{high}$) and 90% of D1R in a low affinity state ($D1^{low}$), while 10% of the D2R are in a low affinity state ($D2^{low}$) and 90% of D2R are in their high affinity state ($D2^{high}$). The overall receptor occupation for each receptor type is then the summed occupation of both states ($D1^{high} + D1^{low}$ and $D2^{high} + D2^{low}$). **(a)** The receptor occupancy at baseline $[DA] = 20nM$ was dominated by the high affinity states for both receptors, even though only 10% of the D1R were in the high state. **(b)** The amount of bound D1R and D2R stayed within the same order of magnitude over a range of baseline [DA]. '$\times$' and '$+$' indicate the model default parameters. **(c)** As in the default model, for a large step up from $[DA] = 20nM$ to $[DA] = 1\mu M$, and **(d)** a step down from $[DA] = 1\mu M$ to $[DA] = 20nM$, D1 and D2 receptor occupancy approached their new equilibrium (EQM, grey dotted lines) only slowly (i.e. over seconds to minutes). As the [D1-DA] changes were dominated by the $D1^{high}$ component, they were very similar to the D2R responses.

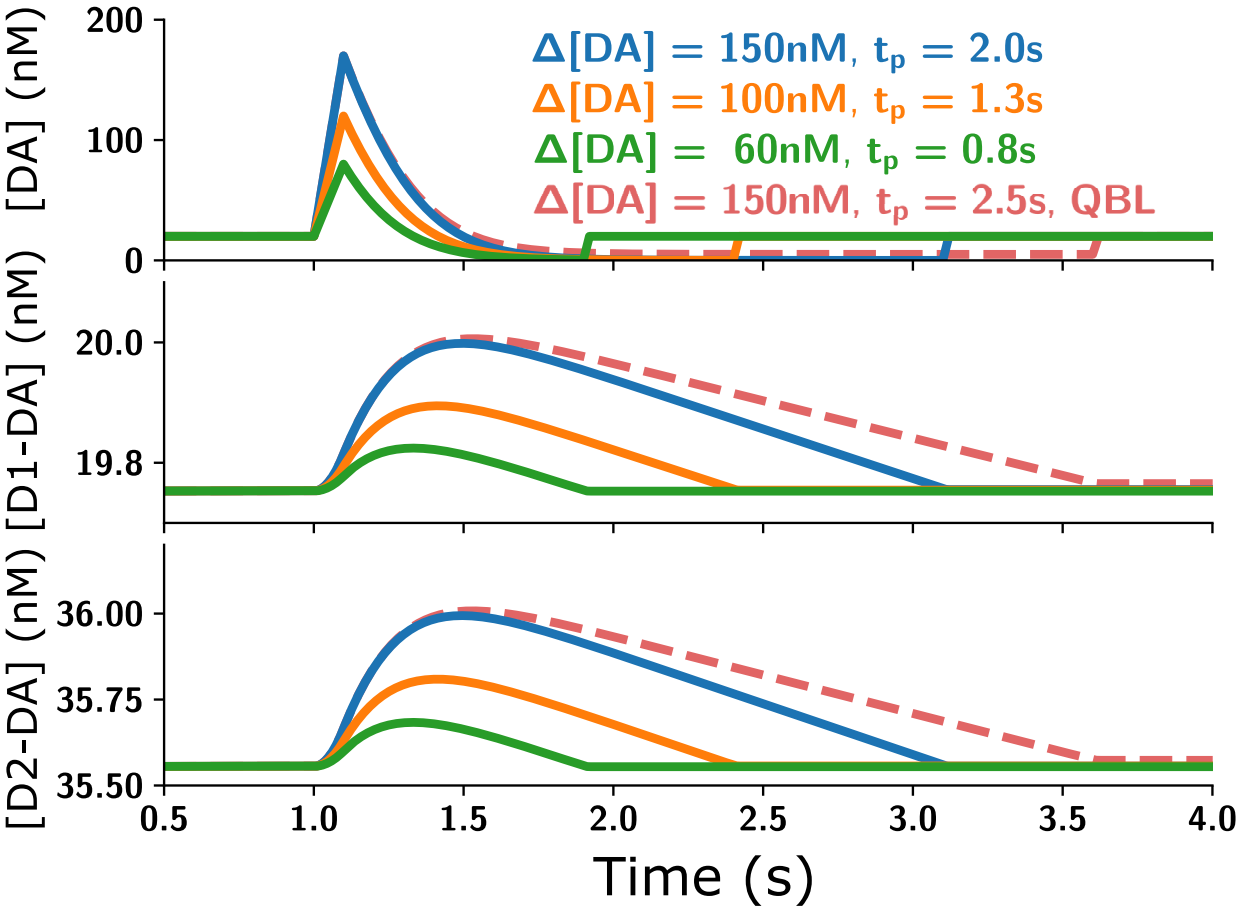
Figure 11: Impact of receptor kinetics on responses to different DA signals with 10% of D1R in a high affinity state ($D1^{high}$) and 10% of D2 receptors in a low affinity state ($D2^{low}$). **(a)** The effect of different phasic DA signals (top panels) on D1 (middle row) and D2 (bottom row) receptor occupancy in the slow kinetics model accounting for affinity states (coloured traces in middle and bottom panels; left scales) and to the affinity-based model (dashed grey traces, right scales). **(b)** Same as in the panel **a** but for DA ramps of different speed. As in the default model, the timing of the maximum receptor occupancy ('$\times$' and 'o' for D1 and D2, respectively) coincides for instant kinetics (purple symbols) with the [DA] peak (combined x and o in top panel), while for slow kinetics (black symbols) it coincides with the offset of the [DA] signal instead (combined '$\times$' and 'o' in top row panel **a**). The main difference to the default model is the higher occupancy of the D1R, due to the $D1^{high}$ component. There is no two-component unbinding since the $D1^{high}$ and $D1^{low}$ have similar off-rates, but differing on-rates. Overall, also for receptors with two affinity states, DA ramps are very effective in occupying the receptors.

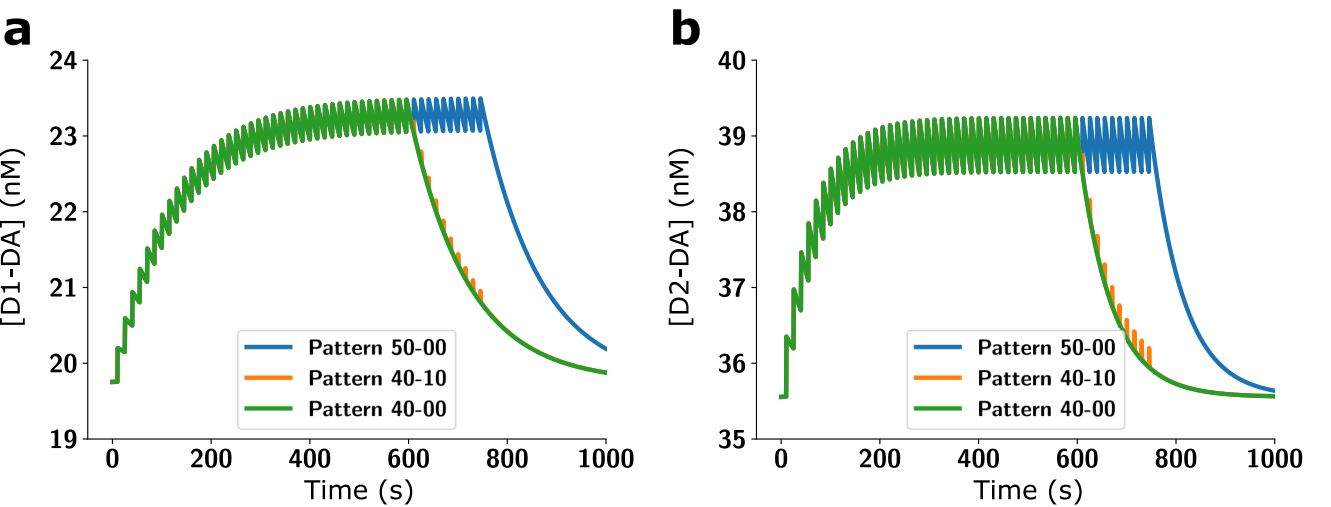
Measured values		
Parameter		Source
$[D1]^m$ in pmol/mg protein	2.840	(Richfield et al., 1989)
$[D2]^m$ in pmol/mg protein	0.696	(Richfield et al., 1989)
ϵ	0.12	(Banay-Schwartz et al., 1992)
α	0.2	(Syková and Nicholson, 2008)
ρ_{brain} in g/ml	1.05	(DiResta et al., 1990)
$f_{D1}^{membrane}$	1.0	(Prou et al., 2001)
$f_{D2}^{membrane}$	0.2	(Prou et al., 2001)
$k_{on}^{D1,orig}$ in $nm^{-1}min^{-1}$	0.00025	(Sano et al., 1979)
$k_{off}^{D1,orig}$ in min^{-1}	0.64	(Sano et al., 1979)
k_{on}^{D2} in $nm^{-1}min^{-1}$	0.02	(Burt et al., 1976)
k_{off}^{D2} in min^{-1}	0.5	(Burt et al., 1976)
Derived Parameters		
Parameter		Source
$[D1]^{tot}$ in nM	≈ 1600	Eq.(17)
$[D2]^{tot}$ in nM	≈ 80	Eq.(17)
$k_{on}^{D1,used}$ in $nm^{-1}min^{-1}$	0.0003125	see Text
$k_{off}^{D1,used}$ in min^{-1}	0.5	see Text

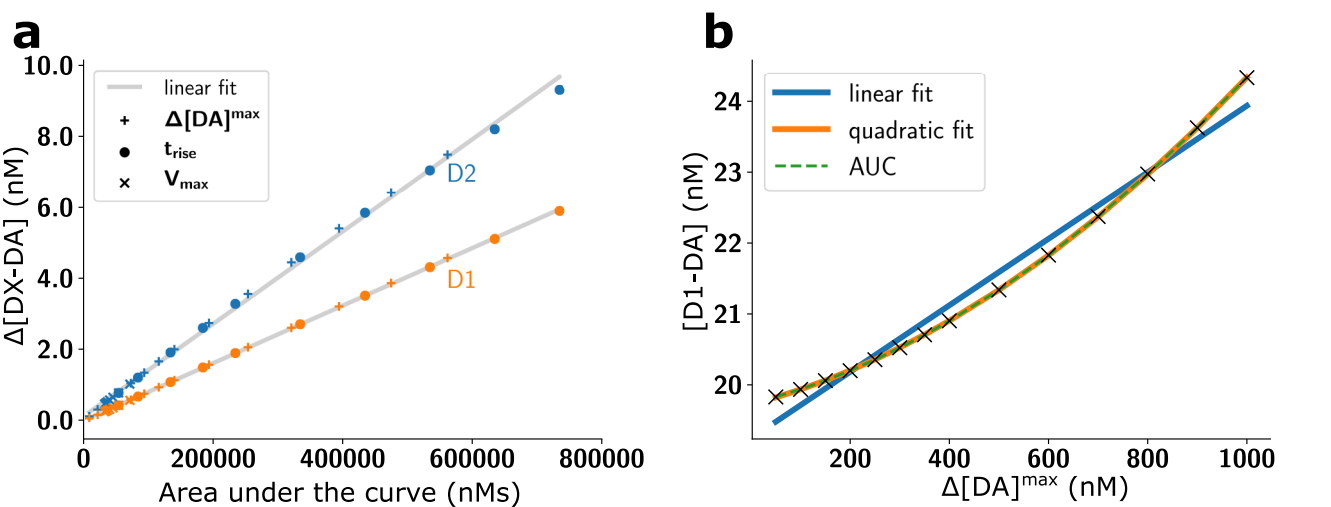
Table 1: Receptor parameters

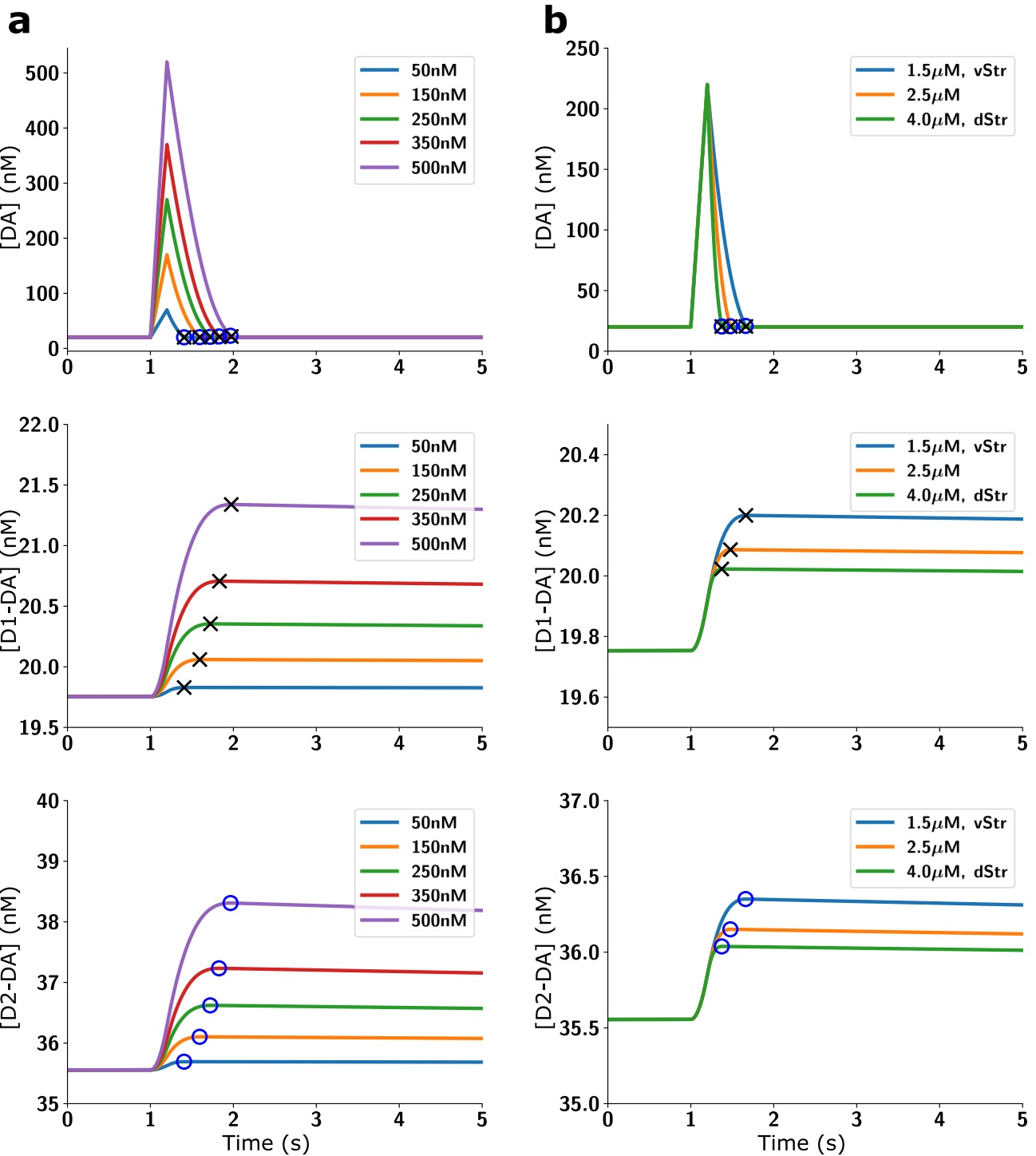

a
K_D^{D1} (μM)
1.0
1.4
1.8
2.2
2.6
3.0
[D1-DA] and [D2-DA] (nM)
0
10
20
30
40
50
60
10
30
50
70
90
110
K_D^{D2} (nM)
b
[D1-DA] and [D2-DA] (nM)
[DA] (nM)
5
10
15
20
25
30
35
40
45
50
c
[D1-DA] and [D2-DA] (nM)
EQM value [D2-DA]
Time (s)
−2
0
2
4
6
8
10
d
[D2-DA] (nM)
[D1-DA] (nM)
EQM value [D2-DA]
EQM value [D1-DA]
Time (s)
0
100
200
300
400
500
600

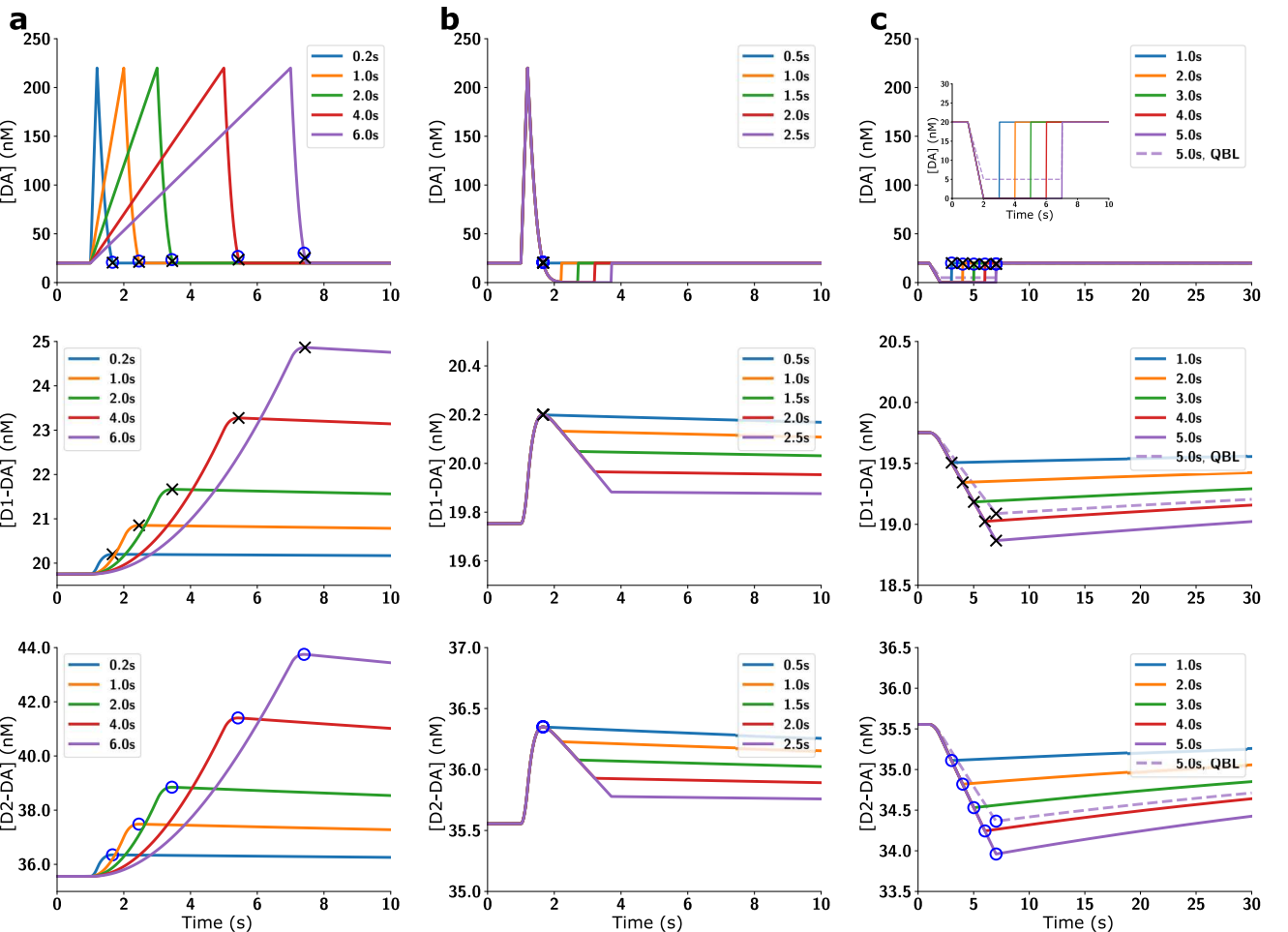

a
b
[DA] (nM)
Long burst
Burst and pause
2s
4s
6s
[D1-DA] (nM)
[D1-DA] (nM)
[D2-DA] (nM)
[D2-DA] (nM)
[DX-DA] (NU)
Long burst D1
Long burst D2
Burst and pause D1
Burst and pause D2
2s D1
2s D2
4s D1
4s D2
6s D1
6s D2
Time (s)
Time (s)

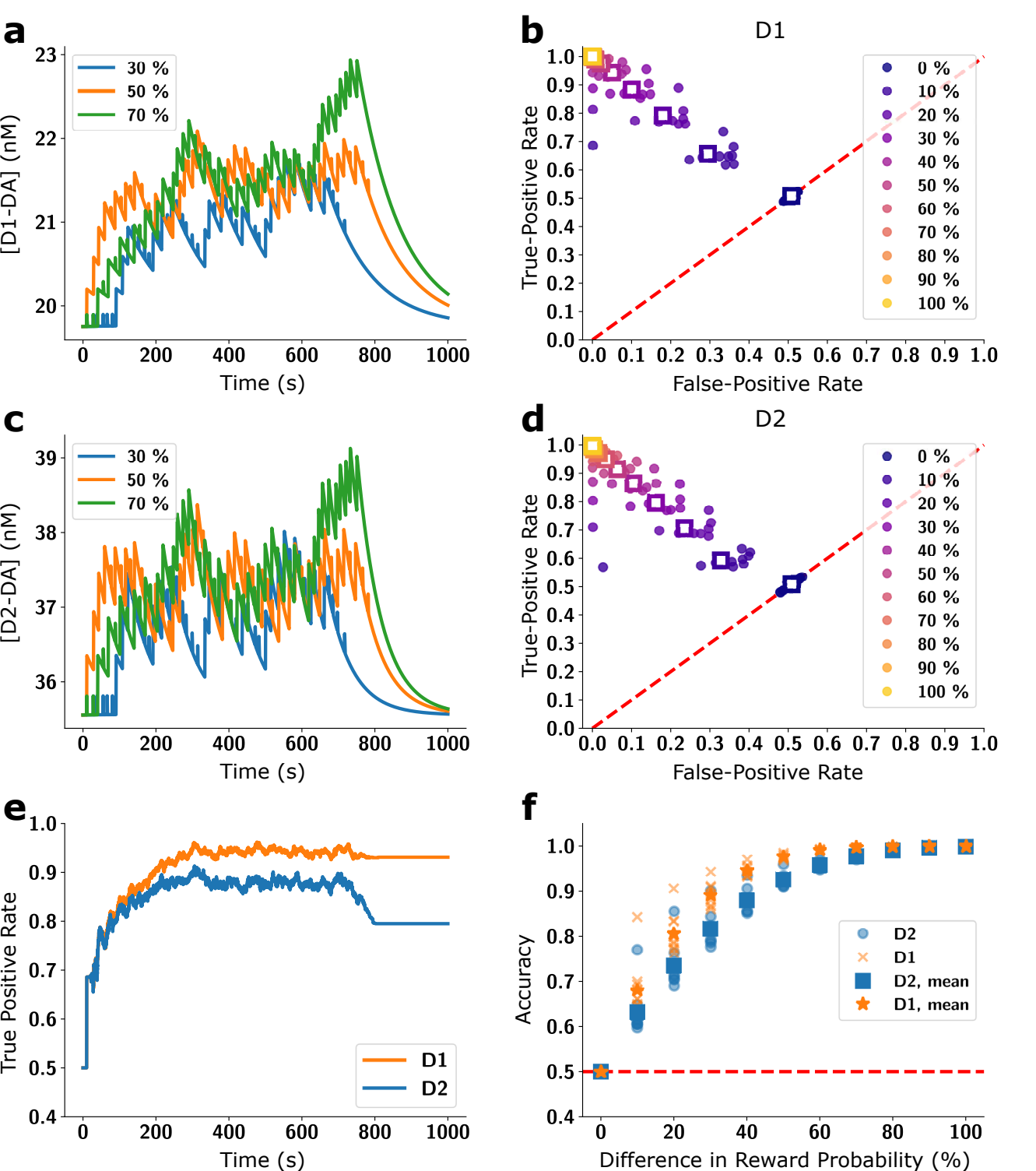

Δ[DA] = 150nM, t_p = 2.0s
Δ[DA] = 100nM, t_p = 1.3s
Δ[DA] = 60nM, t_p = 0.8s
Δ[DA] = 150nM, t_p = 2.5s, QBL
[DA] (nM)
[D1-DA] (nM)
[D2-DA] (nM)
200
100
0
20.0
19.8
36.00
35.75
35.50
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Time (s)

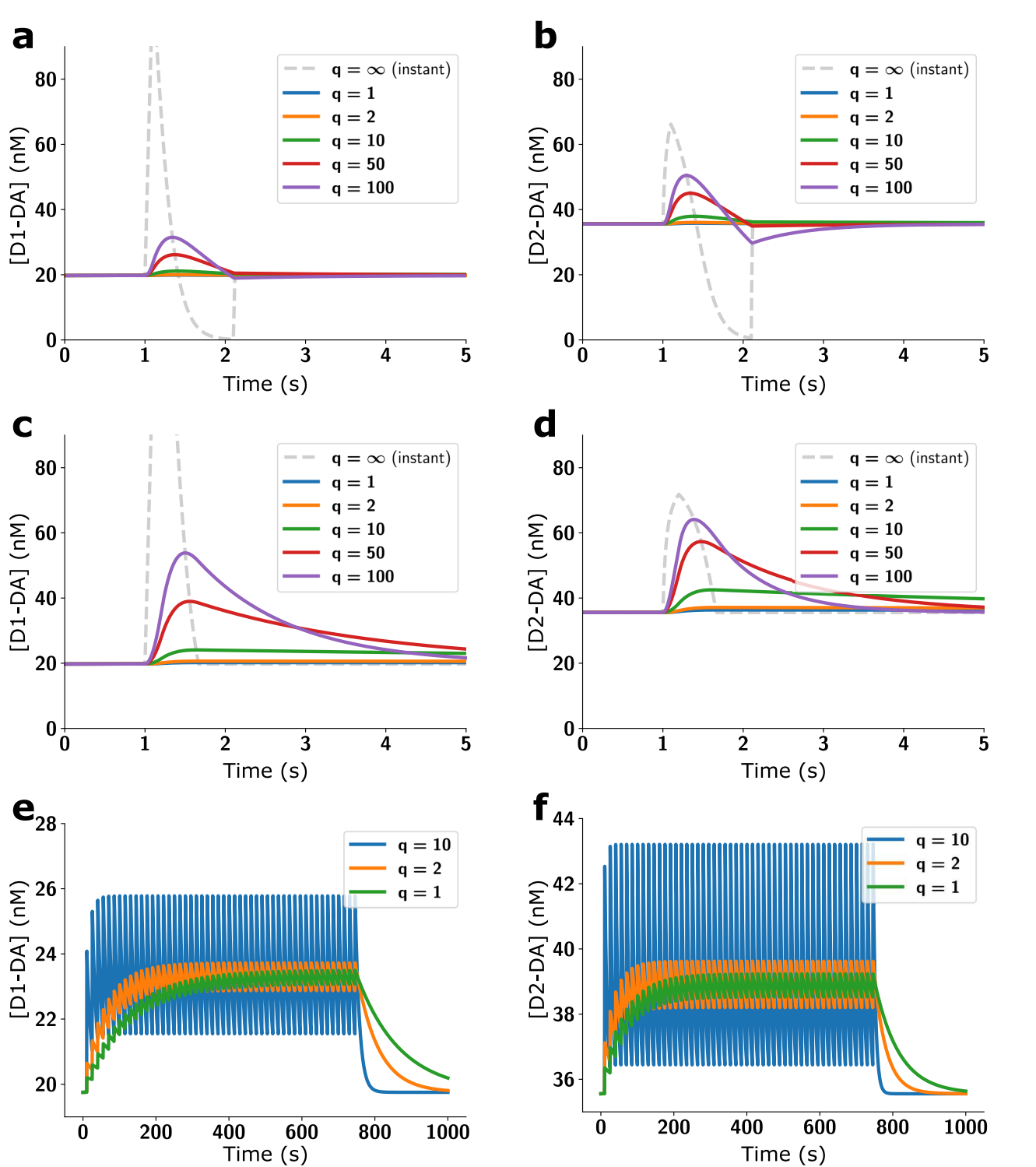

a
b
[D1-DA] (nM)
[D2-DA] (nM)
Time (s)
Time (s)
Pattern 50-00
Pattern 40-10
Pattern 40-00
Pattern 50-00
Pattern 40-10
Pattern 40-00

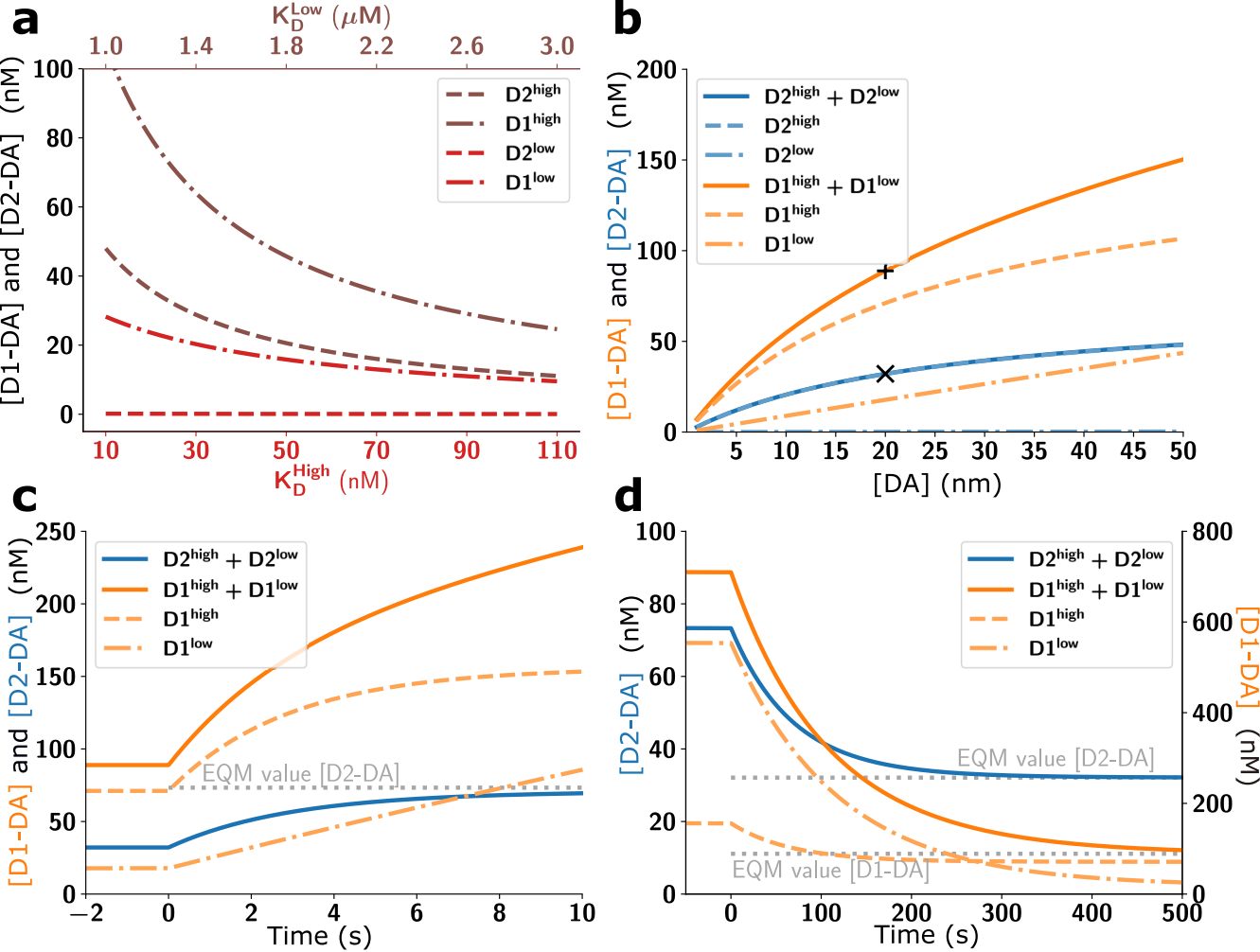

a
linear fit
Δ[DA]max
trise
Vmax
D2
D1
Δ[DX-DA] (nM)
Area under the curve (nMs)
b
linear fit
quadratic fit
AUC
[D1-DA] (nM)
Δ[DA]max (nM)

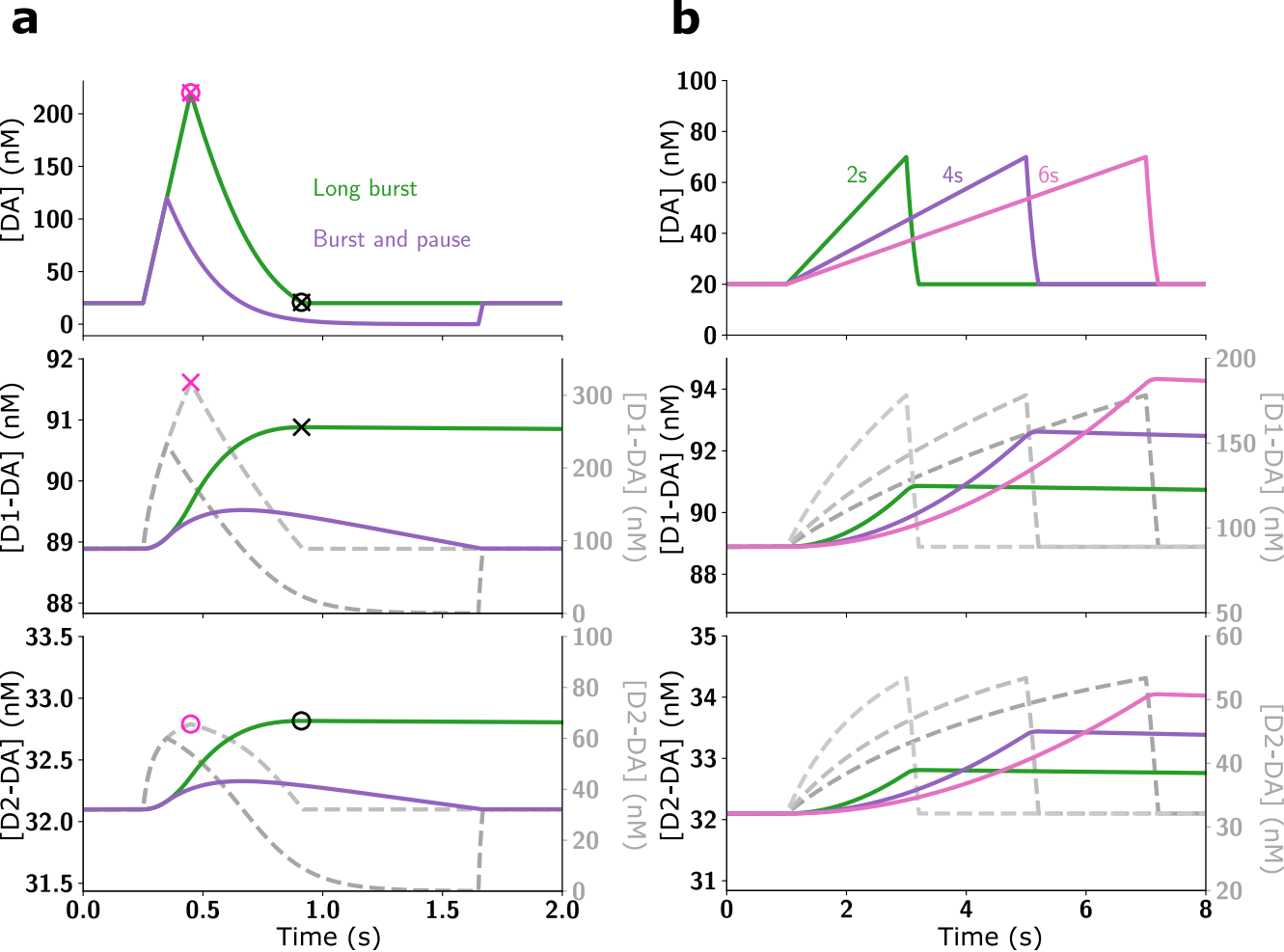

a
b
[DA] (nM)
[D1-DA] (nM)
[D2-DA] (nM)
Time (s)
50nM
150nM
250nM
350nM
500nM
1.5μM, vStr
2.5μM
4.0μM, dStr


a
b
c
[DA] (nM)
[D1-DA] (nM)
[D2-DA] (nM)
Time (s)
0.2s
1.0s
2.0s
4.0s
6.0s
0.5s
1.5s
2.0s
2.5s
3.0s
5.0s
5.0s, QBL


a
30 %
50 %
70 %
[D1-DA] (nM)
Time (s)
b
D1
True-Positive Rate
False-Positive Rate
0 %
10 %
20 %
30 %
40 %
50 %
60 %
70 %
80 %
90 %
100 %
c
[D2-DA] (nM)
d
D2
e
True Positive Rate
D1
D2
f
Accuracy
D2
D1
D2, mean
D1, mean
Difference in Reward Probability (%)


a
b
c
d
e
f
[D1-DA] (nM)
[D2-DA] (nM)
Time (s)
q = ∞ (instant)
q = 1
q = 2
q = 10
q = 50
q = 100


a
K_D^{Low} (μM)
1.0
1.4
1.8
2.2
2.6
3.0
[D1-DA] and [D2-DA] (nM)
100
80
60
40
20
0
D2high
D1high
D2low
D1low
10
30
50
70
90
110
K_D^{High} (nM)
b
[D1-DA] and [D2-DA] (nM)
200
150
100
50
0
D2high + D2low
D2high
D2low
D1high + D1low
D1high
D1low
5
10
15
20
25
30
35
40
45
50
[DA] (nm)
c
[D1-DA] and [D2-DA] (nM)
250
200
150
100
50
0
D2high + D2low
D1high + D1low
D1high
D1low
EQM value [D2-DA]
−2
0
2
4
6
8
10
Time (s)
d
[D2-DA] (nM)
100
80
60
40
20
0
[D1-DA] (nM)
800
600
400
200
0
D2high + D2low
D1high + D1low
D1high
D1low
EQM value [D2-DA]
EQM value [D1-DA]
0
100
200
300
400
500
Time (s)


a
b
[DA] (nM)
[D1-DA] (nM)
[D2-DA] (nM)
Long burst
Burst and pause
2s
4s
6s
Time (s)