Role of CAMKII in reinforcement learning: a computational model of glutamate and dopamine signaling pathways

Shesharao M. Wanjerkhede · Raju S. Bapi

Abstract  Timely release of dopamine (DA) at the striatum seems to be important for reinforcement learning (RL) mediated by the basal ganglia. Houk et al. (in: Houk et al (eds) Models of information processing in the basal ganglia, (1995) proposed a cellular signaling pathway model to characterize the interaction between DA and glutamate pathways that have a role in RL. The model simulation results, using GENESIS KINETIKIT simulator, point out that there is not only prolongation of duration as proposed by Houk et al. (1995), but also an enhancement in the amplitude of autophosphorylation of CaMKII. Further, the autophosphorylated form of CaMKII may form a basis for the “eligibility trace” condition required in RL. This simulation study is the first of its kind to support the comprehensive theoretical proposal of Houk et al. (1995).

Keywords  TDL · Dopamine · Eligibility trace · CaMKII · LTP

1 Introduction

In machine learning literature, recent proposal of mathematical models for learning by reinforcement such as temporal difference learning (TDL) led to successful applications in various domains such as robotics, control, etc. (Sutton and Barto 1998, 1990). TDL framework has also been successfully employed in computational neuroscience for investigation of dopamine (DA) neuron activity (Schultz et al. 1997; Houk et al. 1995). The classical conditioning (stimulus–stimulus association), habit formation (stimulus–action association), and learning of action-outcome contingencies are associated with activity in specific brain regions (Johansen et al. 2009). DA receptors play an important role in such associative learning. At synapses between cortical pyramidal neurons and principal striatal medium spiny neurons (MSNs), postsynaptic D1 and D2 DA receptors are postulated to be necessary for the induction of long-term potentiation (LTP), and depression (LTD), respectively, and such forms of plasticity are thought to underlie associative learning (Shen et al. 2008). Induction of LTP requires activation of NMDA (N-methyl-D-aspartate) receptors by synaptically released glutamate with synergetic postsynaptic membrane depolarization. This relieves the voltage-dependent magnesium block of the NMDA receptor (NMDAR) ion channel allowing calcium (Ca2+) to flow into the dendritic spine (Malenka et al. 1989). Calcium Ca2+ influx through NMDARs is essential for plasticity. The NMDAR-mediated rise in postsynaptic Ca2+ activates a network of signaling molecules that promote persistent changes in synaptic strength, such as LTP (Skeberdis et al. 2006).

Activity of mid-brain DA neurons resembles the reward prediction error signal of the TD model (Schultz et al. 1995; Schultz 1998, 1999; Suri and Schultz 2001). Before learning, the DA neurons are activated by the reward obtained in response to unconditioned stimulus (US). During the learning phase, the activity of these neurons shifts to the reward-predicting stimulus (conditioned stimulus, CS). After learning, the CS predicts the reward and the reward occurs according to the prediction. Thus, prediction error becomes zero. If reward fails to occur after learning, there is a
depression of activity exactly at the predicted time of occurrence of the reward (Schultz 1999). This points to a predictive timing mechanism involved in the DA release at the input of the striatum, the brain area known for its involvement in reinforcement learning (RL) (Schultz 1999; Suri and Schultz 2001; Contreras-Vidal and Schultz 1999). Houk et al. (1995) proposed a cellular signaling model for interaction of DA and glutamate neurotransmitter activities at the striatum that forms the basis for TDL. In this qualitative model, glutamatergic input generates a membrane depolarization through N-methyl-D-aspartate (NMDA), amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor (AMPA), and metabotropic glutamate receptor (mGluR), and opens Ca\(^{2+}\) channels resulting in the influx of Ca\(^{2+}\) into the dendritic spine. This raises the postsynaptic calcium concentration in the dendritic spine and this in turn leads to autophosphorylation of CaMKII. The timely arrival of DA input at the neck of the spinehead generates a cascade of reactions that lead to the prolongation of LTP generated by autophosphorylation of CaMKII (Houk et al. 1995). Since no simulations were done so far to support the qualitative proposal of Houk et al. (1995), the present modeling effort undertook computer simulations in GENESIS/Kinetikit simulation package (Bower and Beeman 1998) to verify this model. Preliminary simulations of the original Houk et al. (1995) model revealed that it is incomplete. Hence, an augmented and detailed signaling model was designed. The redesigned model successfully implements the primary proposal of Houk et al. (1995) that the interaction of DA and glutamatergic pathways leads to prolongation of the duration of autophosphorylation of CaMKII. In addition, it is observed that there is enhancement of the amplitude of autophosphorylation of CaMKII and the model is consistent with several known experimental results and enables us to propose testable predictions.

2 Method

2.1 Dendritic spine model

The model consists of a spherical spinehead and spineneck on the dendrite of a spiny neuron at the striatum as shown in the Fig. 1. In the present model, it is assumed that the glutamate released from the pre-synaptic terminal is located in the external volume of 10 \(\mu\)m\(^3\) surrounding the spinehead (Doi et al. 2005). In the small volume of the spine, the molecules it contains are assumed to be well-mixed and that their concentration is uniform within this volume (Doi et al. 2005). In the model, the volume of the post-synaptic density (PSD) taken is equal 0.002 \(\mu\)m\(^3\), where the NMDARs are located in close proximity with the Ca\(^{2+}\)/calmodulin-dependent protein kinase II (CaMKII). In the simulation, the calcium influx through NMDAR channels is only taken as these channels are prominently involved in phosphorylation of CaMKII, required for LTP (Gesei et al. 1998). Further, only the NR2B subunit of NMDAR is considered since it has been shown that CaMKII directly binds to it (Leonard et al. 1999) and autophosphorylation of CaMKII induces high affinity binding to the NR2B subunit (Leonard et al. 1999; Strack and Colbran 1998). Moreover, it has also been shown that CaMKII phosphorylates NR2B subunit but not NR1 (Omkumar et al. 1996). Glutamate released at the terminals of cortico-striatal afferent binds to NMDAR at the postsynaptic membrane, subsequently the NMDAR channels are opened and Ca\(^{2+}\) enters in to the spinehead. In response to this Ca\(^{2+}\) influx, a cascade of intracellular signaling is initiated. The NMDARs produce a membrane depolarization and an increase in intracellular Ca\(^{2+}\). We did not consider Ca\(^{2+}\)-discharging factors in the model, though the cells have many types of Ca\(^{2+}\)-discharging factors such as Ca\(^{2+}\) binding proteins, Ca\(^{2+}\) pumps, and leaks. Though a detailed model addressing these issues is left for a future study, in the model presented here, we incorporated removal scheme for Ca\(^{2+}\) from spinehead to study the effect on the time course of autophosphorylation of CaMKII. The spinehead diameter, volume etc. are shown in the Table 1.

The spinehead volume for simulation purposes taken is approximately equal to 0.1 \(\mu\)m\(^3\) (Doi et al. 2005; Wilson et al. 1983). The volume of the synaptic cleft at the spine-neck where the DA is released is taken as 3.351 \times 10^{-17} \text{ m}^3 with a diameter equal to 4 \(\mu\)m (Garris et al. 1994). The DA concentration is assumed to be homogeneous within a sphere of 4 \(\mu\)m diameter (Francois 1997). The number of DA molecules contained within the sphere of radius of 2 \(\mu\)m is calculated to be 1,000, and this represents the average number of DA molecules that escape from each synaptic cleft during a stimulus pulse. The number of D1 receptors per synapse is taken to be 1,655 (Garris et al. 1994). With these values for volumes of spherical spinehead and synaptic cleft at the spineneck, the signaling pathways are formulated for basic reactions which are involved in the reward-related learning.

![Fig 1 Spherical spinehead with cortico-striatal and dopaminergic afferents, with spine head diameter equal to 0.53 \(\mu\)m, spine neck diameter equal to 0.3 \(\mu\)m, and spine neck length equal to 2.1 \(\mu\)m](image-url)