Neuropsychopharmacology (2016) 41, 1003–1013

www.neuropsychopharmacology.org

Optogenetic Activation of Adenosine A_{2A} Receptor Signaling in the Dorsomedial Striatopallidal Neurons Suppresses Goal-Directed Behavior

Yan Li[1], Yan He[1], Mozi Chen[1], Zhilan Pu[1], Li Chen[2], Ping Li[3], Bo Li[1], Haiyan Li[1], Zhi-Li Huang[2], Zhihui Li[*,1] and Jiang-Fan Chen[*,1,4]

[1]The Institute of Molecular Medicine, School of Optometry and Ophthalmology and Eye Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China; [2]Department of Pharmacology, Fudan University Shanghai Medical College, Shanghai, China; [3]Molecular Biology Center, Research Institute of Surgery and Daping Hospital, Third Military Medical University, Chongqing, China; [4]Department of Neurology, Boston University, School of Medicine, Boston, MA, USA

The striatum has an essential role in neural control of instrumental behaviors by reinforcement learning. Adenosine A_{2A} receptors (A_{2A}Rs) are highly enriched in the striatopallidal neurons and are implicated in instrumental behavior control. However, the temporal importance of the A_{2A}R signaling in relation to the reward and specific contributions of the striatopallidal A_{2A}Rs in the dorsolateral striatum (DLS) and the dorsomedial striatum (DMS) to the control of instrumental learning are not defined. Here, we addressed temporal relationship and sufficiency of transient activation of optoA_{2A}R signaling precisely at the time of the reward to the control of instrumental learning, using our newly developed *rhodopsin-A_{2A}R chimeras* (*optoA_{2A}R*). We demonstrated that transient light activation of optoA_{2A}R signaling in the striatopallidal neurons in 'time-locked' manner with the reward delivery (but not random optoA_{2A}R activation) was sufficient to change the animal's sensitivity to outcome devaluation without affecting the acquisition or extinction phases of instrumental learning. We further demonstrated that optogenetic activation of striatopallidal A_{2A}R signaling in the DMS suppressed goal-directed behaviors, as focally genetic knockdown of striatopallidal A_{2A}Rs in the DMS enhanced goal-directed behavior by the devaluation test. By contrast, optogenetic activation or focal AAV-Cre-mediated knockdown of striatopallidal A_{2A}R in the DLS had relatively limited effects on instrumental learning. Thus, the striatopallidal A_{2A}R signaling in the DMS exerts inhibitory and predominant control of goal-directed behavior by acting precisely at the time of reward, and may represent a therapeutic target to reverse abnormal habit formation that is associated with compulsive obsessive disorder and drug addiction.
Neuropsychopharmacology (2016) **41,** 1003–1013; doi:10.1038/npp.2015.227; published online 2 September 2015

INTRODUCTION

The striatum has an essential role in neuronal control of the balance between flexible, goal-directed actions and repetitive, habitual behaviors to achieve optimal performance of task (Brown Gould and Graybiel, 2010; Yin and Knowlton, 2006). The striatum is distinguished into the dorsomedial striatum (DMS), which mediates the acquisition and expression of goal-directed behavior through action-outcome learning, and the dorsolateral striatum (DLS), which mediates habit formation through stimulus-response learning (Brown Gould and Graybiel, 2010; Yin and Knowlton, 2006). The shift between goal-directed and habitual actions is associated with changes in neural substrates from DMS to DLS (Yin and Knowlton, 2006) and critically involves the orbitofrontal and striatal circuits (Burguiere *et al*, 2013; Gremel and Costa, 2013). Dysfunction in normal shift between goal-directed and habit actions may contribute to obsessive compulsive disorder (Gillan *et al*, 2011), relapse of drug addiction (Ostlund and Balleine, 2008), habit learning deficit in Parkinson's patients (Knowlton *et al*, 1996), and preservative behaviors of Huntington's disease (Lawrence *et al*, 1998; Redgrave *et al*, 2010). Striatal control of instrumental learning involves critical functions of striatal dopamine and glutamate signaling (Lovinger, 2010; Yin *et al*, 2008): the nigrostriatal dopaminergic pathway provides a 'prediction error' signal for instrumental learning through reinforcement (Rossi *et al*, 2013; Steinberg *et al*, 2013); the activation of glutamatergic corticostriatal pathway is critical to the 'gain' control of cortical incoming information for action-outcome learning (Histed *et al*, 2009; Reynolds *et al*, 2001).

*Correspondence: Professor J-F Chen or Dr Z-H Li, The Institute of Molecular Medicine, School of Optometry and Ophthalmology and Eye Hospital, Wenzhou Medical University, 270 Xueyuan Road, Wenzhou, Zhejiang 325027, China, Tel: +1 781 591 9879, Fax: +86 577 88121352, E-mail: chenjf555@gmail.com or smart_dream2010@yahoo.com.hk
Received 9 April 2015; revised 5 July 2015; accepted 21 July 2015; accepted article preview online 28 July 2015

The adenosine A_{2A} receptors (A_{2A}Rs) are highly enriched in the postsynaptic striatopallidal neurons (Svenningsson *et al*, 1999) where A_{2A}Rs interact with dopamine D_2 receptors (D_2Rs) (Canals *et al*, 2003) and NMDA receptors (Higley and Sabatini, 2010), as well as metabotropic glutamate 5 receptors (Ferre *et al*, 2002). Thus, striatopallidal A_{2A}Rs can integrate incoming information (glutamate) and neuronal sensitivity to this incoming information (dopamine) to control striatal synaptic plasticity and cognitions including goal-directed and habit behaviors (Chen, 2014). Indeed, genetic inactivation of striatal A_{2A}Rs impairs habit formation (Yu *et al*, 2009) and pharmacological reduction of A_{2A}R-mediated cAMP-pCREB signaling in the DMS enhances goal-directed ethanol drinking (Nam *et al*, 2013). However, the contributions of the striatopallidal A_{2A}Rs in the DLS and DMS, two heterogeneous subregions underlying distinct DLS-related habitual or DMS-related goal-directed behavior, to the control of instrumental behavior are not defined.

Furthermore, the reward-based learning mechanism predicts that concurrent activation of the striatal neurons and reward-associated dopaminergic neuron activity is critical to reinforcement learning (Reynolds *et al*, 2001; Schultz *et al*, 1997). However, whether the transient activation of the striatopallidal A_{2A}R signaling precisely at the time of reward is required or sufficient to modify instrumental learning is not known, largely because of the lack of methods to control A_{2A}R signaling in intact animals with required spatiotemporal resolution. To overcome this limitation, we have developed chimeric rhodopsin-A_{2A}R proteins (optoA_{2A}R) by fusing the extracellular and transmembrane domains of rhodopsin with the intracellular loops of the A_{2A}R (Li *et al*, 2015). We leveraged the spatiotemporal resolution of optoA_{2A}R to activate striatopallidal A_{2A}R signaling in a 'time-locked' manner precisely at the time of the reward. Coupling the optoA_{2A}R approach with a satiety-based instrumental learning procedure (Derusso *et al*, 2010), we defined the contribution of striatopallidal A_{2A}R signaling in the DMS and DLS, precisely at or randomly in relation to the time of the reward, to the control of goal-directed and habitual behaviors. We further validated the striatopallidal A_{2A}R control of instrumental learning by focal knockdown of striatopallidal A_{2A}Rs in the DMS and DLS using the AAV-Cre/flox strategy.

MATERIALS AND METHODS

Development of OptoA_{2A}R Strategy

We have developed a optoA_{2A}R, which retains the extracellular and transmembrane domains of rhodopsin (conferring light responsiveness), fused with the intracellular loops of A_{2A}R (conferring specific A_{2A}R signaling), as we described recently (Li *et al*, 2015). The specificity of the optoA_{2A}R signaling was confirmed by light-induced selective enhancement of cAMP and phospho-MAPK levels, by the disappearance of light-induced optoA_{2A}R signaling with a point mutation at the C-terminal region of A_{2A}R, and by the demonstration that optoA_{2A}R activation produced similar activation of signaling, synaptic plasticity, and behavioral responses in intact animals as the A_{2A}R agonist CGS21680 (Li *et al*, 2015). We have constructed viral vectors for optoA_{2A}R (AAV5-EF1α-DIO-mCherry-optoA_{2A}R) and its control (AAV5-EF1α-DIO-mCherry) using a double-floxed inverted (DIO) strategy to target mCherry-optoA_{2A}R fusions in Cre-expressing striatopallidal neurons. The AAV5-EF1α-DIO-mCherry-optoA_{2A}R or AAV5-EF1α-DIO-mCherry was injected to adora2a-cre mice (MMRRC: 031168-UCD) in which the expression of Cre recombinase under the control of A_{2A}R gene regulatory elements was restricted to the striatopallidal neurons (but not cholinergic interneurons or the cortical–striatal projection neurons) (Durieux *et al*, 2009).

Stereotaxic AAV Injection, Optic Fiber Implantation, and Optogenetic Activation of OptoA_{2A}R Signaling

For optoA_{2A}R stimulation experiment, AAV5-EF1α-DIO-mCherry-optoA_{2A}R or AAV5-EF1α-DIO-mCherry (200 nl per striatum) was injected to the DMS (AP, 0.98 mm; ML, 1.20 mm; DV, 2.50 mm) or DLS (AP, 0.98 mm; ML, 2.20 mm; DV, 2.60 mm) of adora2a-cre mice unilaterally. Optic fiber with 200 μm diameter was implanted into relevant brain tissue 0.5 mm above the virus injection site. The mice were maintained for 3 weeks to achieve sufficient virus expression before behavioral training.

Optogenetic stimulation of optoA_{2A}R signaling was achieved by turning on light (473 nm, 10 mW power at the tip) for 2 s per reward (within average 30 or 60 s interval per reward session). To achieve 'time-locked' activation of optoA_{2A}R for 2 s precisely at the time of reward delivery, we programmed optical stimulation to be activated each time contingent on the mouse *active* lever pressing and delivery of sucrose reward (Figure 2b). 'Random' light stimulation was programmed to randomly deliver light in relation to the reward (ie anytime within the interval periods between every two rewards) with same light stimulation parameters as 'time-locked' stimulation (Figure 2b). Light stimulation manipulations were conducted only during random interval (RI) training sessions (Figures 2c, e and 3b).

The Cre-Flox-Mediated Conditional A_{2A}R-Knockdown Strategy

Conditional knockdown of the A_{2A}R gene was achieved by injecting Cre recombinase-expressing AAV into distinct striatal subregions of the A_{2A}R-floxed ($A_{2A}R^{flox/flox}$) mice with the exon 2 of the A_{2A}R gene being flanked by insertion of flox sequences, as we described recently (Lazarus *et al*, 2011). Specifically, AAV8-Cre-zsGreen (200 nl per striatum) was injected into the DMS and DLS of wild-type (WT, $A_{2A}R^{+/+}$) and the floxed ($A_{2A}R^{flox/flox}$) mice bilaterally.

Satiety-Based Instrumental Training

Training session (CRF → RI30 → RI60). Mice were subjected to satiety-based instrumental learning paradigm as we described previously (Yu *et al*, 2009). In brief, mice underwent 3 or 4 days of continuous reinforcement (CRF) training, followed by RI schedule, which promoted habitual behavior: mice were trained 2 days on RI 30 s schedule, followed by 4 days on the RI 60 s schedule (with a 0.1 probability of reward availability every 3 s (RI30) or 6 s (RI60) contingent upon lever pressing).

Devaluation test. Following the training sessions, a 2-day devaluation test was conducted. A specific satiety procedure was applied to alter the current value of a specific reward. On each day, the mice were allowed to have free access to home chows (at least 0.5 g per mouse) or sucrose solution (at least 1 ml per mouse) for at least an hour to achieve sensory-specific satiety. Immediately after the unlimited prefeeding session, mice were given a 5-min extinction test during which the lever was inserted and pressing times was recorded without reward delivery. For each mouse, lever press rate during the devaluation test was normalized to the lever press rate during the last day of RI60 training session before the devaluation test.

Immunofluorescence

Immunofluorescence was performed on free-floating sections (30 μm) using the procedure as we described recently (Augusto *et al*, 2013; Shen *et al*, 2013). Primary antibodies were incubated following the manufacturer's protocols: $A_{2A}R$ (Santa Cruz; 1 : 100), p-MAPK (Cell Signal; 1 : 200), mCherry (Clontech; 1 : 500), enkephalin (Abcam; 1 : 500), and substance-P (Abcam; 1 : 500). Sections were then rinsed and incubated with Alexa 488- or Alexa 594-conjugated secondary antibodies (Invitrogen; 1 : 1000). Slices were washed and mounted and images were acquired and quantified as mean integrated optical density using Image Pro Plus.

Statistical Analysis

Acquisition data were analyzed using two-way ANOVA for repeated measurements with training sessions as within-subjects effect and optoA$_{2A}$R stimulation types or conditional knockdown genotypes as between-subjects effect. For the devaluation test, we performed two-way ANOVA for repeated-measures with optogenetic stimulation types or $A_{2A}R$ conditional knockdown genotypes as one factor and outcome devaluation as another factor. This was followed by simple main-effect analyses to determine the within-subject effect of devaluation test in each group. In addition, as per the experimental design, we also performed planned comparisons within each group between the devalued and valued conditions using a paired *t*-test.

RESULTS

Targeted Expression of OptoA_{2A}R and MAPK Signaling by OptoA_{2A}R Activation in the Striatopallidal Neurons

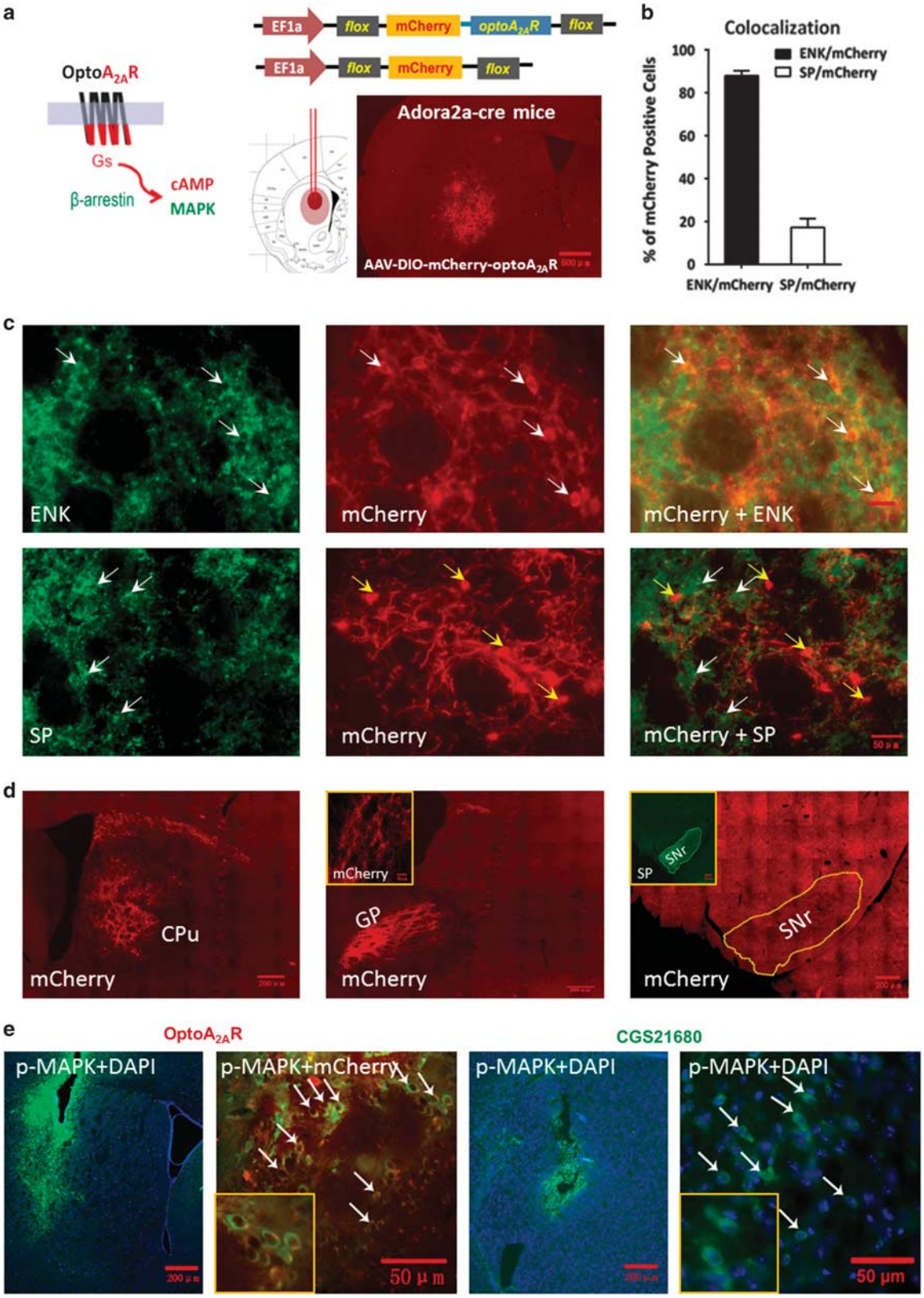
Two weeks after the injection of AAV5-EF1α-DIO-mCherry-optoA$_{2A}$R and its control vector into the striatum of the adora2a-Cre mice (Figure 1a), we verified the selective expression of optoA$_{2A}$R in the striatopallidal neurons. Quantitative analysis of double immunofluorescence staining result indicated that 88% of mCherry (optoA$_{2A}$R-mCherry)-positive cells were colocalized with encephalin (a marker for the striatopallidal neurons), whereas only 17% mCherry-positive cells were colocalized with substance-P (a marker for the striatonigral neurons) in the striatum (Figure 1b). Representative double-immunofluorescence staining images illustrated the colocalization of optoA$_{2A}$R-mCherry with enkephalin but not substance-P (Figure 1c). Furthermore, the red (mCherry) fluorescence was specifically expressed in the terminals of the striatopallidal neurons in the globus pallidus, but was absent in the terminals of striatonigral neurons in the substantia nigra pars reticularta where substance P are highly expressed (Figure 1d). These results confirmed the selective expression of optoA$_{2A}$R in the striatopallidal neurons. Moreover, optoA$_{2A}$R stimulation in the striatum for 5 min induced p-MAPK in the mCherry-positive cells underneath the optic fiber (Figure 1e) in a similar pattern as the $A_{2A}R$ agonist CGS21680. Quantified analysis showed that light-induced p-MAPK activation was detected in 57% mCherry-optoA$_{2A}$R-positive cells ($n = 1218$ from 4 mice). Thus, optoA$_{2A}$R and CGS21680 produced indistinguishable p-MAPK signaling in the striatum.

Optogenetic Activation of Striatopallidal $A_{2A}R$ Signaling in the DMS, Precisely at (but not Randomly in Relation to) the Time of the Reward, Suppressed Goal-Directed Behavior

To determine the effect of optoA$_{2A}$R signaling in the DMS and DLS on goal-directed and habitual actions using a satiety-based instrumental learning paradigm, we first performed an devaluation time-course study to select specific RI training schedule that were most likely sensitive to bidirectional manipulation of the $A_{2A}R$ activity in the DMS and DLS. Devaluation test revealed that after the CFR → RI30 → RI60 training, mice showed a clear goal-directed behavior on the 3rd day, developed habitual behavior on the 4th day, and became a stable habitual behavior on the 5th day after RI60 training (Supplementary Figure 1). Since the mice on the 4th day of RI60 schedule were at the transition period from goal-directed to habitual behavior and were most sensitive to bidirectional manipulation of $A_{2A}Rs$ in the DMS and DLS, we used the RI60 training for 4 days for the rest of the experiments.

We verified that the locations of the optical fiber implantation sites and expression of optoA$_{2A}$R were restricted to the DMS by immunofluorescence (Figure 2a). At the RI sessions, we used the 'time-locked' method to deliver optoA$_{2A}$R stimulation (for 2 s per reward) precisely at the time of reward delivery (Figure 2b). Mice with 'light off' serviced as controls. All mice gradually increased their lever pressing rates to obtain reward and reached the lever pressing plateau at the second day of RI training. There was no main effect of optoA$_{2A}$R stimulation ($F_{1,14} = 0.371$, $p > 0.05$) nor optoA$_{2A}$R stimulation × RI training course interaction effect ($F_{5,70} = 0.098$, $p > 0.05$) by repeated-measures ANOVA. Thus, optogenetic activation of the striatopallidal $A_{2A}R$ signaling in the DMS did neither impair lever pressing performance nor affect acquisition of instrumental learning (Figure 2c).

The devaluation test (Figure 2d) revealed that there was no normalized devaluation × optoA$_{2A}$R interaction effect ($F_{1,14} = 0.429$, $p = 0.523$) by repeated-measures ANOVA. However, preplanned *t*-test showed that the optoA$_{2A}$R mice with 'light off' displayed a goal-directed behavior with sensitivity to devalued reward ($t_{1,7} = 6.861$, $^{***}p < 0.001$, $n = 8$). The goal-directed behavior in the



'light-off' group probably reflects unstable (transient) nature of instrumental behavior for the 4-day RI60 training schedule and might be partially attributed to the relatively low level of lever pressing in this group (and the total rewards received) when the optical fiber implanted in the DMS compared with other experimental groups. Importantly the optoA$_{2A}$R with 'time-locked' stimulation during the RI sessions failed to show sensitivity to outcome devaluation

(preplanned *t*-test, $t_{1,7}=0.709$, $p>0.05$, $n=8$), indicating that their responding was habitual.

To better define the temporal importance of optoA$_{2A}$R signaling precisely at the time of reward and to exclude the nonspecific effect caused by light, we have performed behavioral analyses with separate set of four experimental groups: mice expressing mCherry with 'time-locked' light stimulation ($n=7$), mice expressing optoA$_{2A}$R with 'light off' ($n=9$), mice expressing optoA$_{2A}$R with 'time-locked' light stimulation ($n=8$), and mice expressing optoA$_{2A}$R with 'random' ($n=8$) light stimulation. The light stimulation scheme was illustrated in Figure 2b. Consistent with the result in Figure 2c, there was neither between-subject effect ($F_{3,28}=1.481$, $p=0.241$) nor RI training sessions × manipulation groups interaction effect ($F_{15,140}=1.284$, $p=0.220$) in the acquisition phase by repeated-measures ANOVA (Figure 2e). However, analyses of the devaluation test (Figure 2f) revealed that there was a significant effect of optogenetic manipulation × (normalized) devaluation interaction effect (repeated-measures ANOVA, $F_{3,28}=3.258$, $p=0.036$). The simple main-effect analyses of the devaluation test, respectively, in each group confirmed that only mice with optoA$_{2A}$R expression in the DMS and time-locked light stimulation performed habitually ($F_{1,8}=7.141$, $^{*}p<0.05$ for light off and $F_{1,7}=6.074$, $^{*}p<0.05$ for random stimulation groups, $F_{1,6}=16.050$, $^{**}p<0.01$ for mCherry group). Taken together, statistical analyses of both sets of the experiments (Figure 2d by the preplanned *t*-test and Figure 2f by the repeated-measures ANOVA) support that optogenetic activation of striatopallidal A$_{2A}$R signaling in the DMS modulated the mode of instrumental behaviors by acting precisely at the time of the reward.

Optogenetic Activation of Striatopallidal A_{2A}R Signaling in the DLS had Relatively Limited Effects on Habitual Formation

Next, we examined the effect of optoA$_{2A}$R signaling in the DLS on instrumental behaviors. Similarly, we confirmed the optical fiber implantation sites and expression of optoA$_{2A}$R to be restricted to DLS by immunofluorescence (Figure 3a). Following the RI training sessions, optoA$_{2A}$R mice with 'light off' ($n=10$) or with 'time-locked' stimulation ($n=13$) gradually increased lever presses. There was no main effect of optoA$_{2A}$R stimulation ($F_{1,21}=0.156$, $p>0.05$) and no interaction effect of training session × optoA$_{2A}$R stimulation in the RI sessions ($F_{5,105}=0.916$, $p>0.05$) by repeated-measures ANOVA (Figure 3b). After the 4th day of RI60 training, repeated-measures ANOVA analyses of the devaluation test revealed that there was no optogenetic manipulations × normalized devaluation interaction effect ($F_{1,21}=0.022$, $p=0.884$). However, the preplanned *t*-test showed that optoA$_{2A}$R mice with 'time-locked' stimulation tended to perform goal-directed behavior (*normalized* devaluation test, $t_{1,12}=3.725$, $^{**}p<0.01$ (Figure 3c); devaluation test, $t_{1,12}=2.030$, $p>0.05$ (Supplementary Figure 2c)). Conversely, optoA$_{2A}$R mice with 'light off' displayed habitual behavior (normalized devaluation test, $t_{1,9}=1.270$, $p>0.05$ (Figure 3c); devaluation test, $t_{1,9}=1.868$, $p>0.05$ (Supplementary Figure 2c)). Thus, optogenetic activation of striatopallidal A$_{2A}$R signaling in the DLS tended to promote goal-directed behavior, but its effect was relatively limited.

Knockdown of A_{2A}Rs in the DMS Enhanced Goal-Directed Behavior, Whereas Knockdown of the A_{2A}Rs in the DLS had a Limited Effect on Habitual Behavior

We further evaluated the effects of focal knockdown of the A$_{2A}$Rs in the DMS and DLS on instrumental learning. Figures 4a and 5a provided representative outline of the AAV transfection and A$_{2A}$R focal knockdown areas of the DMS and DLS. Fluorescent images showed that A$_{2A}$Rs expression (the red fluorescence) was reduced selectively in the Cre-expressing regions (indicated by green fluorescence). Quantitative analysis of the A$_{2A}$R immunoreactivity (Figures 4b and 5b) confirmed selective knockdown of A$_{2A}$Rs in the DMS (by 91%) and DLS (by 94%) after transfection with AAV-Cre-zsGreen only in $A_{2A}R^{flox/flox}$ mice but not in WT mice ($A_{2A}R^{+/+}$).

Consistent with the optoA$_{2A}$R results, focal knockdown of A$_{2A}$Rs in the DMS (Figure 4c) and DLS (Figure 5c) did not affect the acquisition of instrumental learning as the $A_{2A}R^{flox/flox}$ and WT mice transfected with AAV-Cre-zsGreen showed identical instrumental learning course at RI training session (DMS: genotype main effect, $F_{1,13}<0.001$, $p>0.05$, RI period × genotype interaction effect: $F_{5,65}=0.859$, $p>0.05$; DLS: genotype main effect, $F_{1,11}=0.534$, $p>0.05$, RI period × genotype interaction effect: $F_{5,55}=1.234$, $p>0.05$; by repeated-measures ANOVA). For the devaluation test, repeated-measures ANOVA analyses revealed that there was genotypes × devaluation interaction effect in the DMS experiment (Figure 4d, normalized devaluation, $F_{1,13}=9.161$, $p=0.01$, simple

Figure 1 Targeted expression and phospho-MAPK (p-MAPK) signaling of optoA$_{2A}$R in striatopallidal neurons. (a) Schematic illustration of the optoA$_{2A}$R chimera construction by replacing the intracellular loops 1, 2, and 3 and C terminal of the bovine rhodopsin with that of the adenosine A$_{2A}$ receptor (A$_{2A}$R) to achieve control of A$_{2A}$R signaling by 473 nm light (left panel). Representative fluorescent image shows the expression of mCherry-optoA$_{2A}$R in the striatum after injection of AAV5-DIO-mCherry-optoA$_{2A}$R to adora2a-cre mice for 2 weeks (right panel). (b) The quantitative data shows that 88% mCherry-positive cells ($n=114$, from four mice) were colocalized with enkephalin (ENK), whereas only 17% mCherry-positive cells ($n=106$, from four mice) were colocalized with substance P (SP). (c) Double immunostaining with the mCherry and the specific antibodies (ENK or SP) showed that optoA$_{2A}$Rs were specifically expressed in ENK-positive striatopallidal neurons (white arrows, upper panels) but not SP-positive striatonigral neurons (yellow arrows, lower panels). (d) Following injection of AAV-DIO-mCherry-optoA$_{2A}$R virus in the dorsomedial striatum (DMS) of adora2a-cre mice, the mCherry fluorescence of striatopallidal projection terminals was specifically expressed in the global pallidum (GP) but not in the substantia nigra pars reticulate (SNr). The green fluorescence of striatonigral projection terminals containing endogenous SP was specifically expressed in the SNr. (e) The expression of p-MAPK was induced by optoA$_{2A}$R stimulation (white arrows, left panels) or CGS21680 injection (white arrows, right panels). Quantified analysis showed that light-induced p-MAPK activation was detected in 57% mCherry-optoA$_{2A}$R-positive cells ($n=1218$ from four mice).

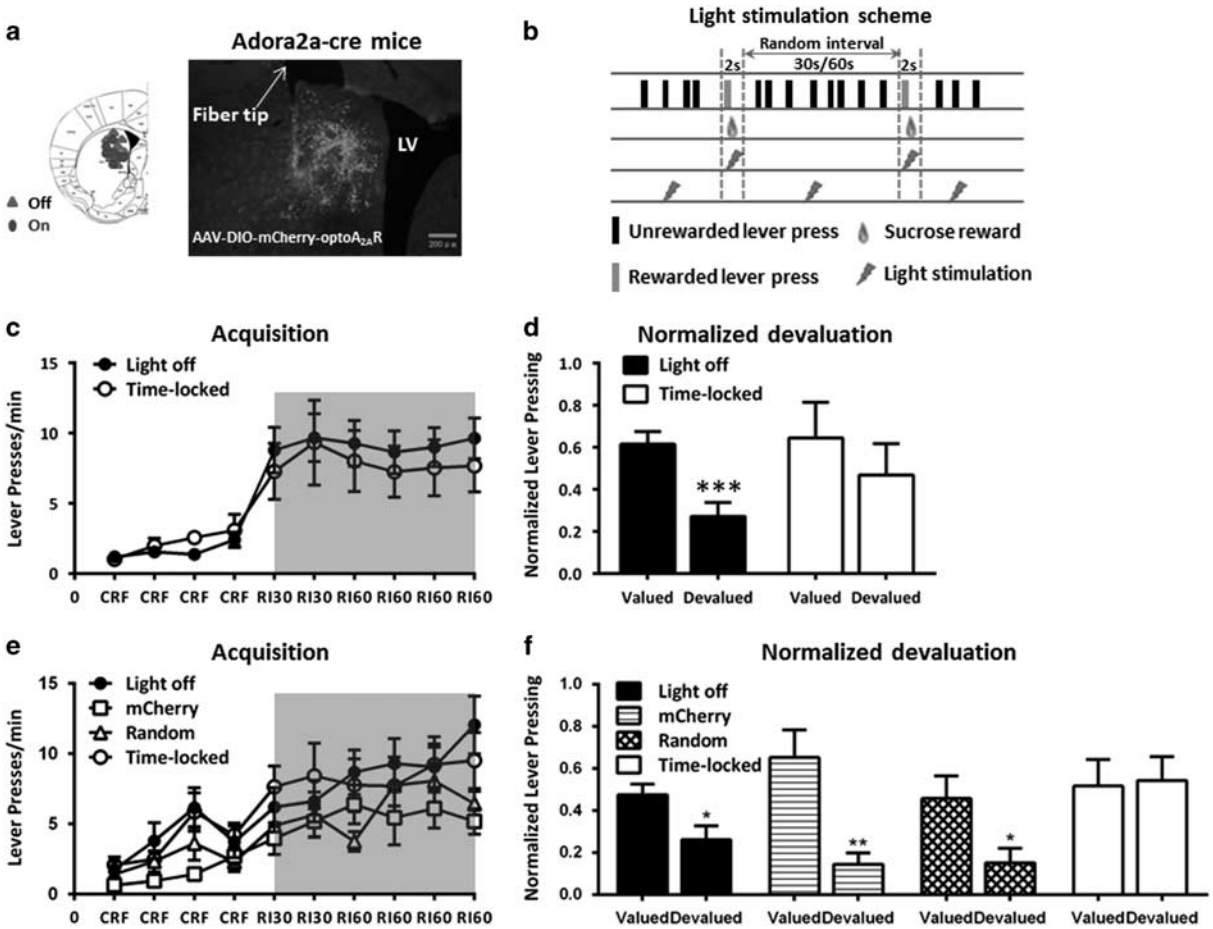


Figure 2 'Time-locked' but not random optogenetic activation of striatopallidal adenosine A_{2A} receptor ($A_{2A}R$) signaling in the dorsomedial striatum (DMS) suppresses goal-directed behavior. (a) Left panel: Schematic illustration of the locations of the fiber tips for each animal in the 'light-off' group (the red triangles) and 'time-locked' activation group (the blue circles). Right panel: Typical coronal section of mCherry-optoA$_{2A}$R expression in the DMS of adora2a-cre(+) mice. The white arrow indicates the optical fiber tip. (b) Schematic illustration of timing of lever pressing, sucrose reward delivery, and optical stimulation. Light stimulation (the blue flash) was delivered to the DMS during a 2-s period in 'time-locked' manner with (the flashes between the two red dotted vertical lines) or in 'random' manner with (the flashes in the random interval periods) reward delivery (the liquid drops). (c) Two groups of mice expressing optoA$_{2A}$R in the DMS were subjected to either 'time-locked' light stimulation or 'light off' ($n=8$ per group) during the random interval (RI) training session (as indicated by the blue bar). The two groups performed indistinguishably in the acquisition phase of instrumental learning by repeated-measures analysis of variance (ANOVA)—RI period × optoA$_{2A}$R stimulation interaction effect: $F_{5,70}=0.098$, $p>0.05$; optoA$_{2A}$R stimulation main effect: $F_{1,14}=0.371$, $p>0.05$. (d) Following the RI training sessions, a 2-day devaluation test without any experimental (optoA$_{2A}$R activation) manipulation was conducted as described in the Materials and Methods section. Mice without optoA$_{2A}$R activation during the RI training sessions significantly reduced their lever presses in devalued condition compared with valued condition (normalized devaluation: $t_{1,7}=6.861$, $^{***}p<0.001$, preplanned *t*-test). By contrast, mice with optoA$_{2A}$R 'time-locked' stimulation showed no significant devaluation effect (normalized devaluation: $t_{1,7}=0.709$, $p>0.05$, preplanned *t*-test). However, there was no normalized devaluation × optoA$_{2A}$R interaction effect by repeated-measures ANOVA analysis ($F_{1,14}=0.429$, $P=0.523$). (e) We further performed instrumental behavioral analyses of a separate set of four experimental groups: mice expressing mCherry with 'time-locked' light stimulation ($n=7$), mice expressing optoA$_{2A}$R with 'light off' ($n=9$), mice expressing optoA$_{2A}$R with 'time-locked' light stimulation ($n=8$), and mice expressing optoA$_{2A}$R with random light stimulation ($n=8$). Consistent with the result in (c) repeated-measures ANOVA analysis indicated that there was neither between-subject effect ($F_{3,28}=1.481$, $p=0.241$) nor RI training sessions × manipulation groups interaction effect ($F_{15,140}=1.284$, $p=0.220$) in the acquisition phase. (f) Repeated-measures ANOVA analyses of the devaluation test revealed that there was significant effect of optogenetic manipulation × (normalized) devaluation interaction effect: $F_{3,28}=3.258$, $p=0.036$. Similarly, the simple main-effect analyses of the devaluation test in four groups indicated that only mice with optoA$_{2A}$R expression in the DMS and time-locked light stimulation performed habitually, whereas other groups displayed goal-directed behavior (simple effect analyses: $F_{1,8}=7.141$, $^{*}p<0.05$ for 'light off' and $F_{1,7}=6.074$, $^{*}p<0.05$ for 'random' stimulation groups, and $F_{1,6}=16.050$, $^{**}p<0.01$ for mCherry group). Data are presented as the mean ± SEM. The color reproduction of this figure is available on the *Neuropsychopharmacology* journal online.

main-effect analyses, $F_{1,6}=35.683$, $^{**}p<0.01$ for $A_{2A}R$ focal knockdown mice; Supplementary Figure 2d, devaluation, $F_{1,13}=10.231$, $p=0.007$, simple main-effect analyses, $F_{1,6}=40.197$, $^{**}p<0.01$ for $A_{2A}R$ focal knockdown mice). This indicated that the control mice displayed a clear habitual action without sensitivity to devaluation condition, whereas focal $A_{2A}R$ knockdown in the DMS altered sensitivity to devaluation by markedly reducing lever presses in the devalued condition. In contrast to the DMS $A_{2A}R$-knockdown effect, focal knockdown of $A_{2A}R$ in the DLS did not affect instrumental behavior and showed no sensitivity to devaluation condition (Figure 5d: genotypes × *normalized* devaluation interaction effect, $F_{1,11}=1.993$, $p=0.186$ by repeated-measures ANOVA, and $t_{1,6}=0.646$,

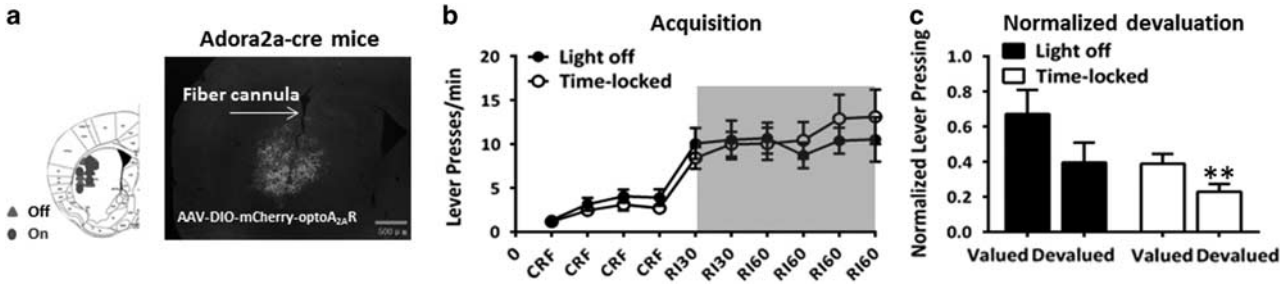


Figure 3 Optogenetic activation of striatopallidal adenosine A_{2A} receptor (A_{2A}R) signaling in the dorsolateral striatum (DLS) exerts relatively limited and possibly opposite control over habitual action compared with the optoA$_{2A}$R in the dorsomedial striatum (DMS). (a) Left: Schematic illustration of the sites of optical fibers implantation. Right: A representative image of mCherry-optoA$_{2A}$R expression and fiber implantation. (b) Mice were under continuous reinforcement (CRF) training followed by RI30 and then RI60 training with or without optoA$_{2A}$R stimulation as described in the Materials and Methods section. The performances of optoA$_{2A}$R mice with 'time-locked' stimulation ($n=13$) or with 'light off' ($n=10$) during the acquisition phase were indistinguishable (repeated-measures analysis of variance (ANOVA), random interval (RI) training course × optogenetic stimulation interaction: $F_{5,105}=0.916$, $p>0.05$; optoA$_{2A}$R stimulation main effect: $F_{1,21}=0.156$, $p>0.05$). (c) OptoA$_{2A}$R mice with 'time-locked' stimulation or 'light off' during the RI training sessions were subjected to devaluation test as described in the Materials and Methods section. Repeated-measures ANOVA analyses revealed that there was no normalized devaluation × optogenetic stimulation interaction effect ($F_{1,21}=0.022$, $p=0.884$). However, preplanned *t*-test analysis revealed that optoA$_{2A}$R mice receiving 'time-locked' stimulation tended to perform goal-directed behavior (only for the *normalized* devaluation test: $t_{1,12}=3.725$, $^{**}p<0.01$; but not for devaluation test: $t_{1,12}=2.030$, $p>0.05$; Supplementary Figure 2c). Whereas optoA$_{2A}$R mice with 'light off' displayed habitual behavior (normalized devaluation test: $t_{1,9}=1.270$, $p>0.05$; devaluation test: $t_{1,9}=1.868$, $p>0.05$; Supplementary Figure 2c). Data are presented as the mean ± SEM. The color reproduction of this figure is available on the *Neuropsychopharmacology* journal online.

$p>0.05$ for DLS A_{2A}R-knockdown mice, $t_{1,5}=2.017$, $p>0.05$ for WT mice by preplanned *t*-test; the devaluation test showed a similar result; Supplementary Figure 2e). Thus, consistent with the results of the optoA$_{2A}$R, these findings validate that focal knockdown of striatopallidal A_{2A}Rs in the DMS selectively enhanced goal-directed behavior, whereas focal knockdown of striatopallidal A_{2A}Rs in the DLS had little effect on habitual behavior.

DISCUSSION

Transient and 'Time-Locked' Activation of optoA$_{2A}$R Signaling Precisely at the Time of Reward is Required and Sufficient to Modulate Goal-Directed Behavior

The contemporary theory of striatum-dependent learning postulates that the concurrent activation of presynaptic nigral–striatal dopamine (reinforcement) signaling and corticostriatal glutamate (sensorimotor) signaling and post-synaptic striatopallidal neuronal activity (modulated by neuromodulator such as adenosine) is critical to striatal synaptic plasticity and instrumental learning (Yagishita *et al*, 2014; Reynolds *et al*, 2001; Schultz *et al*, 1997). Indeed, modification of instrumental learning by optogenetic manipulation of striatal neurons was only effective in a narrow temporal window (ie before or concurrent with the onset of cue (Tai *et al*, 2012), or in the time segment (1.5 s) between action selection and outcome (Aquili *et al*, 2014)), supporting the temporal importance of dopamine, glutamate, and neuromodulator signaling in striatum-dependent instrumental learning. Different from rapid neurotransmitter release such as dopamine and glutamate, extracellular adenosine is generated by conversion of ATP to adenosine through a set of ectonucleotidases and by bidirectional nucleotide transporters (Chen *et al*, 2013). Striatopallidal A_{2A}R activity may modulate instrumental learning by acting precisely at the time of the reward to integrate dopamine or glutamate signaling for coding the action-outcome contingency. Alternatively, striatopallidal A_{2A}Rs control instrumental learning by modulating the vigor of actions (Desmurget and Turner, 2010), by providing permissive role in learning association (Brainard and Doupe, 2000), or by modulating the 'off-line' processing of incoming signaling (glutamate) (Pomata *et al*, 2008). In these alternative schemes, the temporal relationship between striatopallidal activity (ie A_{2A}R activity) and the reward is not essential. Thus, a critical question is whether the transient activation of A_{2A}R precisely at the time of reward delivery was required and sufficient to modulate instrumental learning. This question has not been addressed owing to the lack of methods to control A_{2A}R signaling in behaving animals with required temporal resolution. Our development of the optoA$_{2A}$R (Li *et al*, 2015) offers the opportunity to optogenetically control the A_{2A}R signaling with sufficient temporal resolution. We showed that transient (2 s per reward) and 'time-locked' light activation of the optoA$_{2A}$R signaling in the striatopallidal neurons precisely at the time of the reward (but not random light stimulation) was required and sufficient to modify the sensitivity to outcome devaluation without affecting the acquisition. The requirement and sufficiency of 'time-locked' and transient activation of optoA$_{2A}$R signaling at the time of the reward to modify instrumental learning demonstrated a temporally specific relationship between adenosine A_{2A}R signaling and nigrostriatal dopamine signaling in association with the reward delivery and possibly corticostriatal glutamate signaling that converged on the striatopallidal neurons. Considering the extensive interaction between A_{2A}Rs, D_2Rs, and NMDA receptors in the striatopallidal neurons (Lovinger, 2010), we speculate that concurrent activation of A_{2A}Rs, D_2Rs, and NMDA receptors in the striatopallidal neurons allows the integration of adenosine, dopamine, and glutamate signaling, and coding of the mode of instrumental learning behavior (Abeliovich *et al*, 1992; Tai *et al*, 2012).

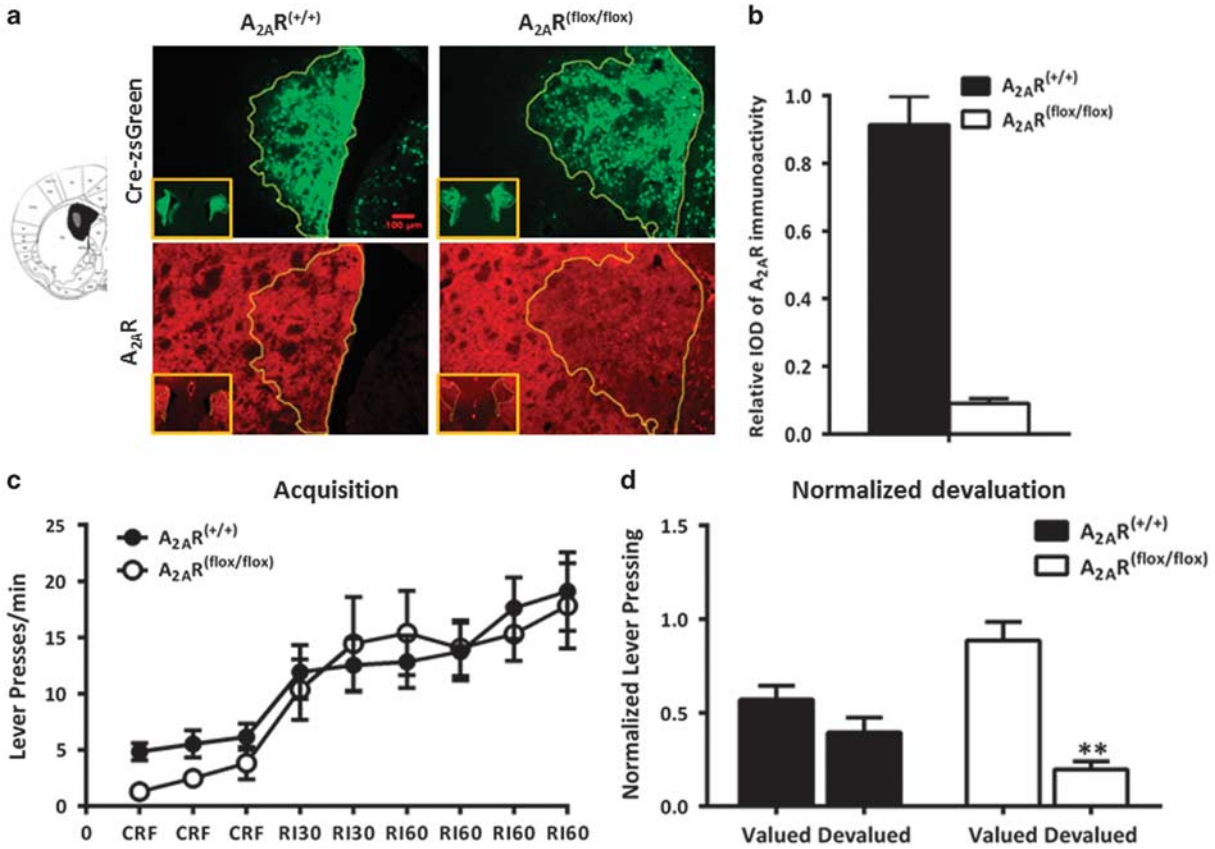


Figure 4 Focal knockdown of adenosine A_{2A} receptors (A_{2A}Rs) in the dorsomedial striatum (DMS) enhances goal-directed behavior. (a) Left: Schematic illustration of the maximal (black) and minimal (gray) A_{2A}R knockdown areas in the DMS. Right: Representative immunofluorescent photomicrographs show focal knockdown expression of A_{2A}Rs in the DMS after injection of AAV-Cre-zsGreen into the $A_{2A}R^{(flox/flox)}$ (right panels) and $A_{2A}R^{(+/+)}$ mice (left panels). Intensity of A_{2A}Rs (red) were significantly deceased in the overlapping area with zsGreen expression (the yellow circle) in $A_{2A}R^{(flox/flox)}$ mice but not in $A_{2A}R^{(+/+)}$ mice. (b) Quantitative analysis showed that A_{2A}R expression were markedly reduced in the virus-transfected regions of $A_{2A}R^{(flox/flox)}$ mice compared with $A_{2A}R^{(+/+)}$ mice. (c) Two–three weeks after bilateral injection of AAV-Cre-zsGreen into the DMS, $A_{2A}R^{(flox/flox)}$ mice and $A_{2A}R^{(+/+)}$ mice ($n=8$ per group) were under CRF-RI30-RI60 training paradigm as described in the Materials and Methods section. Both groups similarly increased their lever pressing rate during the acquisition phases (repeated-measures analysis of variance (ANOVA) revealed no random interval (RI) period × genotype interaction effect: $F_{5,65}=0.859$, $p>0.05$; and no genotype main effect: $F_{1,13}<0.001$, $p>0.05$). (d) Mice with DMS A_{2A}R knockdown significantly reduced their lever pressing in the devalued condition compared with that of the valued condition, but the $A_{2A}R^{(+/+)}$ mice responded insensitively to the selective satiety devaluation treatment (normalized devaluation × genotype interaction effect: $F_{1,13}=9.161$, $p=0.01$; simple effect analysis: $F_{1,6}=35.683$, $^{**}p<0.01$ for A_{2A}R focal knockdown mice by repeated-measures ANOVA). Data are presented as the mean ± SEM. CRF, continuous reinforcement. The color reproduction of this figure is available on the *Neuropsychopharmacology* journal online.

The Striatopallidal A_{2A}R Signaling in the DMS Provides a 'Break' Mechanism to Constrain Instrumental Learning

As the DMS and DLS are distinctly involved in goal-directed and habitual behaviors, respectively (Balleine *et al*, 2009; Brown Gould and Graybiel, 2010; Yin and Knowlton, 2006), another important question is whether striatopallidal A_{2A}Rs exert DMS- and DLS-specific control over instrumental learning. Our bidirectional manipulation of the striatopallidal A_{2A}Rs by optogenetic activation of A_{2A}R signaling and Cre-mediated knockdown of A_{2A}Rs in the DMS and DLS demonstrated that A_{2A}Rs in the DMS exerted an inhibitory and predominant control of goal-directed, whereas striatopallidal A_{2A}Rs in the DLS had relatively limited but possibly opposite effects on habit formation. This is consistent with the associative corticostriatal–DMS loop being 'default' model of striatal function (Thorn *et al*, 2010) and with previous finding that deletion of the indirect pathway in the DMS (but not DLS) produces pronounced psychomotor and cognitive effects (Durieux *et al*, 2012). This view is also supported by recent pharmacological study that reduction of A_{2A}R-mediated PKA-pCREB signaling in the DMS enhanced acquisition of goal-directed ethanol drinking behaviors in mice (Nam *et al*, 2013). Given the prominent role of the DMS in control of goal-directed behavior, our finding that focal knockdown of striatopallidal A_{2A}Rs in the DMS captures the goal-directed characteristics of striatum-specific A_{2A}R knockout (KO) mice argue that striatum-A_{2A}R KO mice displayed enhanced goal-directed behavior, but manifested as impaired habit formation (Yu *et al*, 2009). Although our analysis is designed to isolate the striatopallidal A_{2A}R action from other action sites, this does not preclude the contribution of the A_{2A}Rs in extrastriatal or cholinergic neurons to the control of instrumental learning, which needed to be further defined.

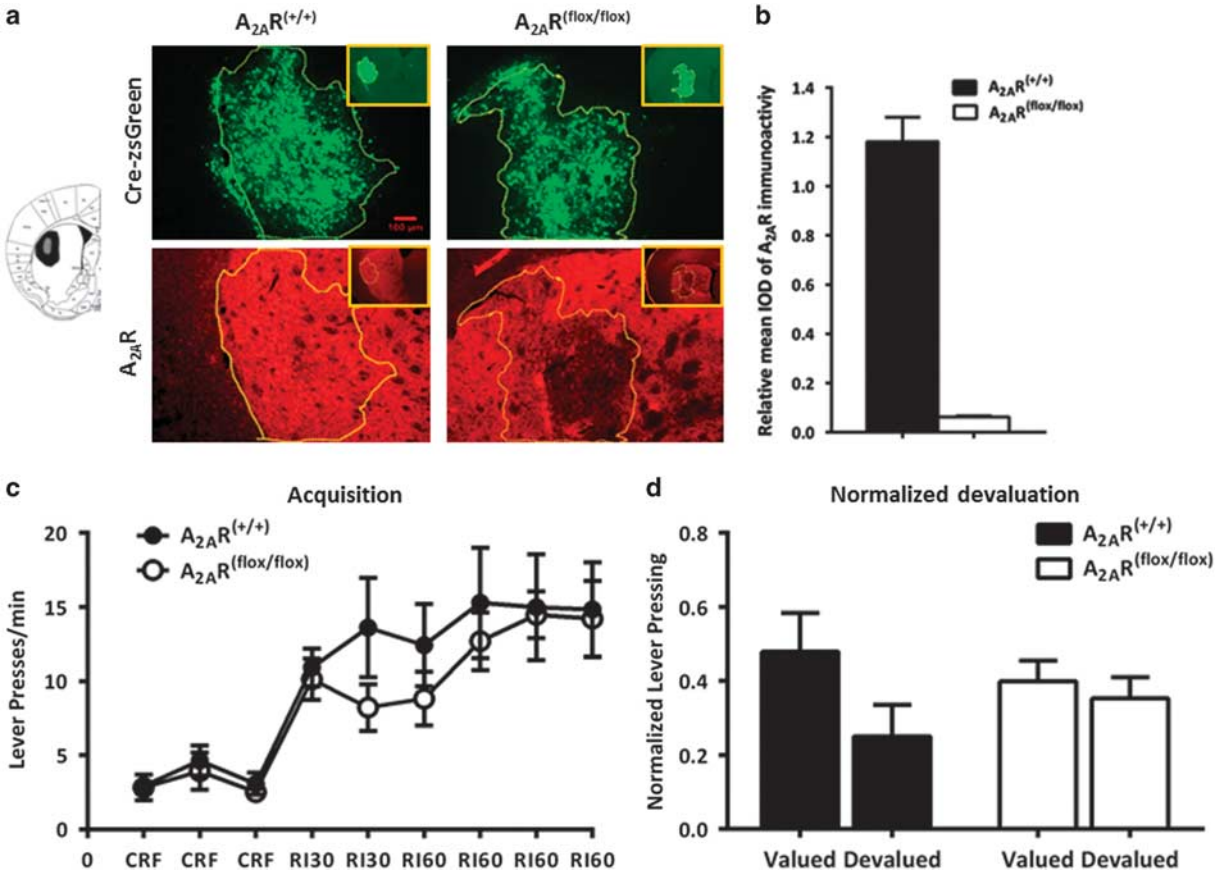


Figure 5 Focal knockdown of the adenosine A_{2A} receptors (A_{2A}Rs) in the dorsolateral striatum (DLS) exerts relatively limited effects on habitual behaviors. (a) Representative image shows that A_{2A}Rs were knocked down selectively in the area with the AAV-Cre-zsGreen expression in the DLS of $A_{2A}R^{(flox/flox)}$ mice but not in $A_{2A}R^{(+/+)}$ mice. The yellow circle depicted the boundary of the AAV-Cre-zsGreen expression and A_{2A}R knockdown area. (b) Quantitative analysis shows that A_{2A}R expression was markedly reduced in the AAV-Cre-zsGreen transfected regions of $A_{2A}R^{(flox/flox)}$ mice compared with $A_{2A}R^{(+/+)}$ mice. (c) Focal A_{2A}R knockdown in the DLS ($n=7$) did not affect lever pressing during the acquisition phase compared with their $A_{2A}R^{(+/+)}$ controls ($n=6$) (repeated-measures analysis of variance (ANOVA) revealed no random interval (RI) period × genotype interaction effect: $F_{5,55}=1.234$, $p>0.05$; and no genotype main effect: $F_{1,11}=0.534$, $p>0.05$). (d) There was no genotype × devaluation interaction effect ($F_{1,11}=1.993$, $p=0.186$, repeated-measures ANOVA) for the normalized devaluation test. Both groups similarly showed insensitivity to outcome devaluation (DLS A_{2A}R knockdown mice: normalized devaluation; $t_{1,6}=0.646$, $p>0.05$; wild-type (WT) mice: normalized devaluation; $t_{1,5}=2.017$, $p>0.05$). Data are presented as the mean ± SEM. CRF, continuous reinforcement.

It is worth noting that similar to striatal A_{2A}R KO (Yu *et al*, 2009), either optoA_{2A}R activation or focal A_{2A}R knockdown of striatopallidal A_{2A}R activity did not affect the acquisition (Figures 2c, 3b, 4c and 5c) or omission/extinction (Supplementary Figure 3) phase of instrumental learning, but specifically affect sensitivity to outcome devaluation. The lack of the optoA_{2A}R effect during the acquisition and extinction/omission phases indicates that striatopallidal A_{2A}Rs unlikely affect general arousal status or attention to influence instrumental learning, but instead it may modify the motivation control of action selection. This notion is consistent with the critical role of striatopallidal A_{2A}Rs in the modulation of effort expenditure and motivation (Mingote *et al*, 2008; Nunes *et al*, 2013).

Lastly, bidirectional manipulation of the striatopallidal A_{2A}Rs by optoA_{2A}R and Cre-mediated A_{2A}R knockdown demonstrates a critical role of the postsynaptic striatopallidal A_{2A}Rs and the striatopallidal pathway in the DMS in control of instrumental learning. This collaborates with the recent finding that transient optogenetic stimulation of striatopallidal neurons introduces opposing biases during decision making in mice (Tai *et al*, 2012), and that loss of striatal long-term depression largely restricted to striatopallidal neurons is associated with a shift in behavioral control from goal-directed action to habitual responding (Nazzaro *et al*, 2012). Taken together with increasing evidences from diverse learning paradigms that striatopallidal A_{2A}Rs assume an inhibitory control over working memory (Wei *et al*, 2011; Zhou *et al*, 2009), fear condition (Singer *et al*, 2013; Wei *et al*, 2014), reversal learning (Wei *et al*, 2011), and instrumental learning (Yu *et al*, 2009), we postulate that postsynaptic striatopallidal A_{2A}R function may provide a 'break' mechanism to constrain some cognitions including instrumental learning (Chen, 2014). If the postulated 'break' mechanism of the striatopallidal A_{2A}R is validated by future experiments, this provides a framework for a pharmacological strategy by blocking striatopallidal A_{2A}R activity to reverse abnormal

habit formation that is associated with compulsive obsessive disorder and relapse of drug addiction.

FUNDING AND DISCLOSURE

This study was sponsored by the Start-up Fund from Wenzhou Medical University (No. 89211010 JFC; No. 89212012, JFC and KYQD121004, ZL), the Zhejiang Provincial Special Funds (No. 604161241), the Special Fund for Building National Clinical Key Resource (Key Laboratory of Vision Science, Ministry of Health, No. 601041241), the Central Government Special Fund for Local Universities' Development (No. 474091314), the Zhejiang Provincial Natural Science Foundation of China (No. LQ15H090007), and by NIH Grants (NS041083-11 and NS073947) and special BUSM research fund DTD 4-30-14. The authors have no proprietary interest in any materials or methods described within this article.

ACKNOWLEDGMENTS

We thank Liu Ya for her assistance in image analysis of optical fluorescent density.

REFERENCES

Abeliovich A, Gerber D, Tanaka O, Katsuki M, Graybiel AM, Tonegawa S (1992). On somatic recombination in the central nervous system of transgenic mice. *Science* **257**: 404–410.

Aquili L, Liu AW, Shindou M, Shindou T, Wickens JR (2014). Behavioral flexibility is increased by optogenetic inhibition of neurons in the nucleus accumbens shell during specific time segments. *Learn Mem* **21**: 223–231.

Augusto E, Matos M, Sevigny J, El-Tayeb A, Bynoe MS, Muller CE *et al* (2013). Ecto-5'-nucleotidase (CD73)-mediated formation of adenosine is critical for the striatal adenosine A2A receptor functions. *J Neurosci* **33**: 11390–11399.

Balleine BW, Liljeholm M, Ostlund SB (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res* **199**: 43–52.

Brainard MS, Doupe AJ (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* **404**: 762–766.

Brown Gould B, Graybiel AM (2010). Afferents to the cerebellar cortex in the cat: evidence for an intrinsic pathway leading from the deep nuclei to the cortex. 1976. *Cerebellum* **9**: 1–13.

Burguiere E, Monteiro P, Feng G, Graybiel AM (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* **340**: 1243–1246.

Canals M, Marcellino D, Fanelli F, Ciruela F, de Benedetti P, Goldberg SR *et al* (2003). Adenosine A2A-dopamine D2 receptor–receptor heteromerization: qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J Biol Chem* **278**: 46741–46749.

Chen JF (2014). Adenosine receptor control of cognition in normal and disease. *Int Rev Neurobiol* **119**: 257–307.

Chen JF, Eltzschig HK, Fredholm BB (2013). Adenosine receptors as drug targets—what are the challenges? *Nat Rev Drug Discov* **12**: 265–286.

Derusso AL, Fan D, Gupta J, Shelest O, Costa RM, Yin HH (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front Integr Neurosci* **4**: 17 (doi: 10.3389/fnint.2010.00017).

Desmurget M, Turner RS (2010). Motor sequences and the basal ganglia: kinematics, not habits. *J Neuroscience* **30**: 7685–7690.

Durieux PF, Bearzatto B, Guiducci S, Buch T, Waisman A, Zoli M *et al* (2009). D2R striatopallidal neurons inhibit both locomotor and drug reward processes. *Nat Neurosci* **12**: 393–395.

Durieux PF, Schiffmann SN, de Kerchove d'Exaerde A (2012). Differential regulation of motor control and response to dopaminergic drugs by D1R and D2R neurons in distinct dorsal striatum subregions. *EMBO J* **31**: 640–653.

Ferre S, Karcz-Kubicha M, Hope BT, Popoli P, Burgueno J, Gutierrez MA *et al* (2002). Synergistic interaction between adenosine A2A and glutamate mGlu5 receptors: implications for striatal neuronal function. *Proc Natl Acad Sci USA* **99**: 11940–11945.

Gillan CM, Papmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW *et al* (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am J Psychiatry* **168**: 718–726.

Gremel CM, Costa RM (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* **4**: 2264.

Higley MJ, Sabatini BL (2010). Competitive regulation of synaptic Ca^{2+} influx by D2 dopamine and A2A adenosine receptors. *Nat Neurosci* **13**: 958–966.

Histed MH, Pasupathy A, Miller EK (2009). Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron* **63**: 244–253.

Knowlton BJ, Mangels JA, Squire LR (1996). A neostriatal habit learning system in humans. *Science* **273**: 1399–1402.

Lawrence AD, Sahakian BJ, Robbins TW (1998). Cognitive functions and corticostriatal circuits: insights from Huntington's disease. *Trends Cogn Sci* **2**: 379–388.

Lazarus M, Shen HY, Cherasse Y, Qu WM, Huang ZL, Bass CE *et al* (2011). Arousal effect of caffeine depends on adenosine A2A receptors in the shell of the nucleus accumbens. *J Neurosci* **31**: 10067–10075.

Li P, Rial D, Canas PM, Yoo JH, Li W, Zhou X *et al* (2015). Optogenetic activation of intracellular adenosine A receptor signaling in the hippocampus is sufficient to trigger CREB phosphorylation and impair memory. *Mol Psychiatry*, 1–11 (doi:10.1038/mp.2014.182).

Lovinger DM (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* **58**: 951–961.

Mingote S, Font L, Farrar AM, Vontell R, Worden LT, Stopper CM *et al* (2008). Nucleus accumbens adenosine A2A receptors regulate exertion of effort by acting on the ventral striatopallidal pathway. *J Neurosci* **28**: 9037–9046.

Nam HW, Hinton DJ, Kang NY, Kim T, Lee MR, Oliveros A *et al* (2013). Adenosine transporter ENT1 regulates the acquisition of goal-directed behavior and ethanol drinking through A2A receptor in the dorsomedial striatum. *J Neurosci* **33**: 4329–4338.

Nazzaro C, Greco B, Cerovic M, Baxter P, Rubino T, Trusel M *et al* (2012). SK channel modulation rescues striatal plasticity and control over habit in cannabinoid tolerance. *Nat Neurosci* **15**: 284–293.

Nunes EJ, Randall PA, Podurgiel S, Correa M, Salamone JD (2013). Nucleus accumbens neurotransmission and effort-related choice behavior in food motivation: effects of drugs acting on dopamine, adenosine, and muscarinic acetylcholine receptors. *Neurosci Biobehav Rev* **37**(Part A): 2015–2025.

Ostlund SB, Balleine BW (2008). On habits and addiction: an associative analysis of compulsive drug seeking. *Drug Discov Today Dis Models* **5**: 235–245.

Pomata PE, Belluscio MA, Riquelme LA, Murer MG (2008). NMDA receptor gating of information flow through the striatum *in vivo*. *J Neurosci* **28**: 13384–13389.

Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H *et al* (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat Rev Neurosci* **11**: 760–772.

Reynolds JN, Hyland BI, Wickens JR (2001). A cellular mechanism of reward-related learning. *Nature* **413**: 67–70.

Rossi MA, Sukharnikova T, Hayrapetyan VY, Yang L, Yin HH (2013). Operant self-stimulation of dopamine neurons in the substantia nigra. *PLoS One* **8**: e65799.

Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* **275**: 1593–1599.

Shen HY, Canas PM, Garcia-Sanz P, Lan JQ, Boison D, Moratalla R *et al* (2013). Adenosine A(2)A receptors in striatal glutamatergic terminals and GABAergic neurons oppositely modulate psychostimulant action and DARPP-32 phosphorylation. *PLoS One* **8**: e80902.

Singer P, Wei CJ, Chen JF, Boison D, Yee BK (2013). Deletion of striatal adenosine A(2A) receptor spares latent inhibition and prepulse inhibition but impairs active avoidance learning. *Behav Brain Res* **242**: 54–61.

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* **16**: 966–973.

Svenningsson P, Le Moine C, Fisone G, Fredholm BB (1999). Distribution, biochemistry and function of striatal adenosine A2A receptors. *Progr Neurobiol* **59**: 355–396.

Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* **15**: 1281–1289.

Thorn CA, Atallah H, Howe M, Graybiel AM (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* **66**: 781–795.

Wei CJ, Augusto E, Gomes CA, Singer P, Wang Y, Boison D *et al* (2014). Regulation of fear responses by striatal and extrastriatal adenosine A2A receptors in forebrain. *Biol Psychiatry* **75**: 855–863.

Wei CJ, Singer P, Coelho J, Boison D, Feldon J, Yee BK *et al* (2011). Selective inactivation of adenosine A(2A) receptors in striatal neurons enhances working memory and reversal learning. *Learn Mem* **18**: 459–474.

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

Yin HH, Knowlton BJ (2006). The role of the basal ganglia in habit formation. *Nat Rev Neurosci* **7**: 464–476.

Yin HH, Ostlund SB, Balleine BW (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur J Neurosci* **28**: 1437–1448.

Yu C, Gupta J, Chen JF, Yin HH (2009). Genetic deletion of A2A adenosine receptors in the striatum selectively impairs habit formation. *J Neurosci* **29**: 15100–15103.

Zhou SJ, Zhu ME, Shu D, Du XP, Song XH, Wang XT *et al* (2009). Preferential enhancement of working memory in mice lacking adenosine A(2A) receptors. *Brain Res* **1303**: 74–83.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (http://www.nature.com/npp)