



# Pharmacological Blockade of Adenosine A<sub>2A</sub> but Not A<sub>1</sub> Receptors Enhances Goal-Directed Valuation in Satiety-Based Instrumental Behavior

Yan Li<sup>1†</sup>, Xinran Pan<sup>2†</sup>, Yan He<sup>2†</sup>, Yang Ruan<sup>2</sup>, Linshan Huang<sup>2</sup>, Yuling Zhou<sup>2</sup>, Zhidong Hou<sup>2</sup>, Chaoxiang He<sup>2</sup>, Zhe Wang<sup>1</sup>, Xiong Zhang<sup>1\*</sup> and Jiang-Fan Chen<sup>2,3\*</sup>

<sup>1</sup> Department of Neurology, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, <sup>2</sup> School of Optometry and Ophthalmology and Eye Hospital, The Institute of Molecular Medicine, Wenzhou Medical University, Wenzhou, China, <sup>3</sup> Department of Neurology, School of Medicine, Boston University, Boston, MA, United States

## OPEN ACCESS

### Edited by:

Francisco Ciruela,  
Universitat de Barcelona, Spain

### Reviewed by:

Elena Martín-García,  
Universitat Pompeu Fabra, Spain  
Sebastiano Alfio Torrisi,  
Università degli Studi di Catania, Italy

### \*Correspondence:

Xiong Zhang  
zhangxiong98@gmail.com  
Jiang-Fan Chen  
chenjf@bu.edu

† These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Experimental Pharmacology and Drug  
Discovery,  
a section of the journal  
Frontiers in Pharmacology

Received: 06 December 2017

Accepted: 05 April 2018

Published: 24 April 2018

### Citation:

Li Y, Pan X, He Y, Ruan Y, Huang L,  
Zhou Y, Hou Z, He C, Wang Z,  
Zhang X and Chen J-F (2018)  
Pharmacological Blockade  
of Adenosine A<sub>2A</sub> but Not A<sub>1</sub>  
Receptors Enhances Goal-Directed  
Valuation in Satiety-Based  
Instrumental Behavior.  
Front. Pharmacol. 9:393.  
doi: 10.3389/fphar.2018.00393

The balance and smooth shift between flexible, goal-directed behaviors and repetitive, habitual actions are critical to optimal performance of behavioral tasks. The striatum plays an essential role in control of goal-directed versus habitual behaviors through a rich interplay of the numerous neurotransmitters and neuromodulators to modify the input, processing and output functions of the striatum. The adenosine receptors (namely A<sub>2A</sub>R and A<sub>1</sub>R), with their high expression pattern in the striatum and abilities to interact and integrate dopamine, glutamate and cannabinoid signals in the striatum, may represent novel therapeutic targets for modulating instrumental behavior. In this study, we examined the effects of pharmacological blockade of the A<sub>2A</sub>Rs and A<sub>1</sub>Rs on goal-directed versus habitual behaviors in different information processing phases of instrumental learning using a satiety-based instrumental behavior procedure. We found that A<sub>2A</sub>R antagonist acts at the coding, consolidation and expression phases of instrumental learning to modulate animals' sensitivity to goal-directed valuation without modifying action-outcome contingency. However, pharmacological blockade and genetic knockout of A<sub>1</sub>Rs did not affect acquisition or sensitivity to goal-valuation of instrumental behavior. These findings provide pharmacological evidence for a potential therapeutic strategy to control abnormal instrumental behaviors associated with drug addiction and obsessive-compulsive disorder by targeting the A<sub>2A</sub>R.

**Keywords:** adenosine A<sub>2A</sub> receptor, adenosine A<sub>1</sub> receptor, goal-directed behavior, habit, instrumental behavior

## INTRODUCTION

Goal-directed and habitual behaviors are crucial adaptive behaviors for our daily life. Goal-directed behavior evaluates actions prospectively and can flexibly adjust action depending on environmental changes, but this comes at the cost of more cognitive resource. By contrast, habitual behavior is usually developed after repeated overtraining for days and represents automatic responses elicited by external or internal triggers during the performance of routine procedures with less cognitive loads (Dolan and Dayan, 2013). These two behavioral processes can develop in parallel or sequentially and can also reciprocally compete with each other for behavioral control

(Yin and Knowlton, 2006; Balleine and O'Doherty, 2010; Kim and Hikosaka, 2015). The balance between flexible goal-directed actions and repetitive habitual behaviors has an essential role in achieving optimal performance of behavioral task. Dysregulation of goal-directed versus habitual behaviors is considered to be a potential mechanism underlying the relapse of drug addiction (Ostlund and Balleine, 2008), obsessive compulsive disorder (Gillan et al., 2011; Robbins et al., 2012; Burguiere et al., 2015), and may contribute to the executive dysfunction in Parkinson's (Redgrave et al., 2010; de Wit et al., 2011) and Huntington's disease patients (Lawrence et al., 1998).

The striatum plays an essential role in control of goal-directed versus habitual behaviors (Yin and Knowlton, 2006; Graybiel and Grafton, 2015; Kim and Hikosaka, 2015). The dorsal medial striatum (DMS)-connecting orbitofrontal cortex (OFC) is critical for goal-directed valuation (Gremel and Costa, 2013), while the dorsal lateral striatum (DLS) and its connecting infralimbic cortex act as dual operators for habitual behavioral control (Smith and Graybiel, 2013a,b). Additionally, the accumbens nucleus (NAc)-ventral Pallidum (VP) pathway is necessary for goal-directed valuation as inactivation of NAc-VP pathway impairs the predictive learning (Leung and Balleine, 2013). Furthermore, the nigro-striatal dopamine signaling acts as a prediction error and motivational signal to drive instrumental learning (Glimcher, 2011; Rossi et al., 2013; Steinberg et al., 2013). Thus, the striatum acts as a key locus in integrating the cortico-striatal glutamate and the substantia nigra-striatal dopamine signals to control goal-directed and habitual behaviors.

The striatal control of instrumental behaviors is accomplished through a rich interplay of the numerous neurotransmitters and neuromodulators to modify the input, processing and output functions of the striatum (Lovinger, 2010). Several studies have documented the involvement of the D<sub>2</sub> receptor (Kwak et al., 2014), cannabinoid receptor type 1 (CB<sub>1</sub>R) (Hilario et al., 2007) and 5-hydroxytryptamine 6 (5-HT<sub>6</sub>) receptor (Eskenazi et al., 2015) in control of instrumental behavior. However, pharmacological control of instrumental behaviors is under-explored and the effective pharmacological strategies for the control of goal-directed versus habitual behaviors are lacking. Adenosine A<sub>1</sub> and A<sub>2A</sub> receptors are highly expressed in the striatum and are increasingly recognized as important pharmacological targets for controlling cognition under normal and disease conditions (Chen et al., 2013; Chen, 2014). The Gs-coupled facilitating A<sub>2A</sub> receptor (A<sub>2A</sub>R) and Gi-coupled inhibitory A<sub>1</sub> receptor (A<sub>1</sub>R) both integrate dopamine (Shen W. et al., 2008), glutamate (Kreitzer and Malenka, 2007), and BDNF (Tebano et al., 2008; Wei et al., 2014) signaling to modulate synaptic plasticity and control cognition. For example, using our newly developed chimeric rhodopsin-A<sub>2A</sub>R proteins (optoA<sub>2A</sub>R), we recently demonstrated that transient activation of A<sub>2A</sub>R by light in a time-locked manner with reward delivery is sufficient to impair goal-directed behavior whereas focal knockdown of A<sub>2A</sub>R in the striatum enhances goal-directed behaviors (Yu et al., 2009; Li et al., 2016). Similarly, pharmacological blockade of A<sub>2A</sub>R promoted goal-directed seeking for ethanol in ENT1 knockout mice (Nam et al., 2013b) and restored goal-directed sensitivity to negative feedback

in the methamphetamine (METH)-paired context (Furlong et al., 2017). These pharmacological, genetic, and optogenetic demonstrations of the cognitive "brake" mechanism of A<sub>2A</sub>R activation led us to propose that pharmacological blockade of the A<sub>2A</sub>R represents a promising therapeutic target for controlling goal-directed behaviors.

As the first step in developing an adenosine receptor-based pharmacological approach to control the goal-directed versus habitual behaviors, we coupled the A<sub>2A</sub>R antagonist (KW6002) and A<sub>1</sub>R antagonist (DPCPX) with the satiety-based instrumental learning paradigm to address the effect of pharmacological blockade of the A<sub>2A</sub>R and A<sub>1</sub>R on three aspects of instrumental learning processes: (i) behavioral elements of instrumental behaviors (i.e., acquisition of action-outcome contingency versus goal-evaluation) by acquisition of instrumental behavior, the devaluation test and the omission test; (ii) the instrumental learning processes by administering the A<sub>2A</sub>R antagonist either prior to the training (learning/encoding) or post-training (consolidation) during the random interval (RI) schedule, or immediately before the devaluation and omission tests (expression/retrieval of instrumental behaviors); (iii) the potential role of the A<sub>1</sub> receptor in control of instrumental learning.

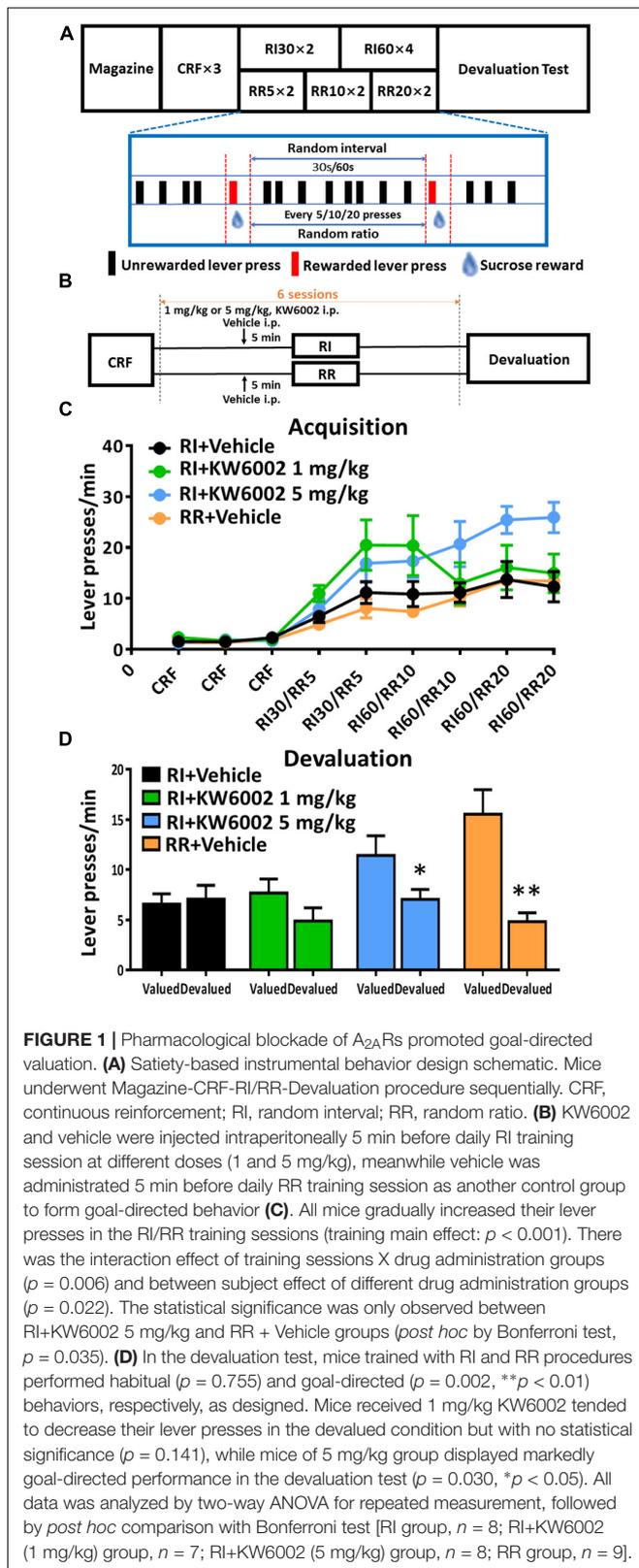
## MATERIALS AND METHODS

### Animals

Animals were handled in accordance with the protocols approved by the Institutional Ethics Committee for Animal Use in Research and Education at Wenzhou Medical University, China. C57BL/6 male mice at least 8 weeks old (23–27 g each) were used in the experiments. The A<sub>1</sub>R knockout mice (A<sub>1</sub>R<sup>-/-</sup>=+/+) and wild-type littermate controls (A<sub>1</sub>RC=C) have been well characterized previously (Johansson et al., 2001) and confirmed by PCR analysis of gene identification before the experiment. Mice were housed in an ambient temperature of 22 ± 0.5°C and a relative humidity of 60 ± 2% with a 12 h light/dark cycle. Mice were single-housed and underwent experiments in the light cycle.

### Satiety-Based Instrumental Training and Testing

All instrumental learning experiments were performed in standard operant chambers (Med Associates). Each chamber was equipped with a retractable lever on either side of a pump with a syringe that delivered liquid reward (20% sucrose solution, 20 μl/reinforce which can be suspended from the syringe) and a house light (3 W, 24 V) mounted on the opposite side of the chamber. Training and testing procedures were performed following Rossi et al (Rossi and Yin, 2012) and illustrated in **Figure 1A**. In brief, mice were first given one 30-min magazine training session during which the sucrose solution was delivered on a random time 60 s schedule with the lever removed. Three days of continuous reinforcement (CRF) training sessions were followed to sufficiently establish the initial association between lever press and reward. At the start of the session, the house light was illuminated, and one lever was inserted into



the chamber. The house light remained illuminated and the lever remained inserted and active during the entire session.

During CRF session, each lever press resulted in the delivery of one drop of 20  $\mu$ l 20% sucrose solution. Sessions ended after 60 min or when 50 rewards had been earned, whichever came first. After CRF, mice underwent RI schedule which was critical for habitual learning. They were trained 2 days on RI 30 s, with a 0.1 probability of reward availability every 3 s contingent upon lever pressing, followed by 4 days on the 60 s interval schedules (0.1 probability of reward availability every 6 s contingent upon lever pressing). Just as CRF training, RI sessions ended after 60 min or when 50 rewards had been earned, whichever came first. To further confirm goal-directed behavioral pattern, we also employed random ratio (RR) training paradigm as control which contributed to goal-directed behavior. Progressively leaner schedules of reinforcement were used: CRF for 3 days, then RR 5 for 2 days (RR5; each response was rewarded at a probability of 0.2 on average), RR10 for 2 days and finally RR20 for 2 days. In the training sessions, home chows were given 1.5–2g daily to maintain 80–85% of their free-feeding weight.

Following the RI/RR 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, which were used for maintaining their weights in the training sessions or sucrose solution which was earned by their lever pressing for at least an hour to achieve sensory-specific satiety. Immediately after the unlimited pre-feeding session, mice were given a 5-min extinction test during which the lever was inserted and pressing times were recorded without reward delivery. The order of the valued and devalued condition tests (day 1 or day 2) was counterbalanced across animals. Mice sensitive to manipulation of outcome value would significantly reduce their lever presses on the devalued condition compared with the valued condition. Then after two supplementary RI60 training sessions, mice were further evaluated by a 30-min omission test in which action-outcome contingency was altered. In the omission test, mice had to control their lever-press impulsion formed by previous training sessions for 20 s to obtain the reward. Any lever press would reset the time counter and mice would hold another 20 s not to press the lever for reward delivery.

## Drug Administration

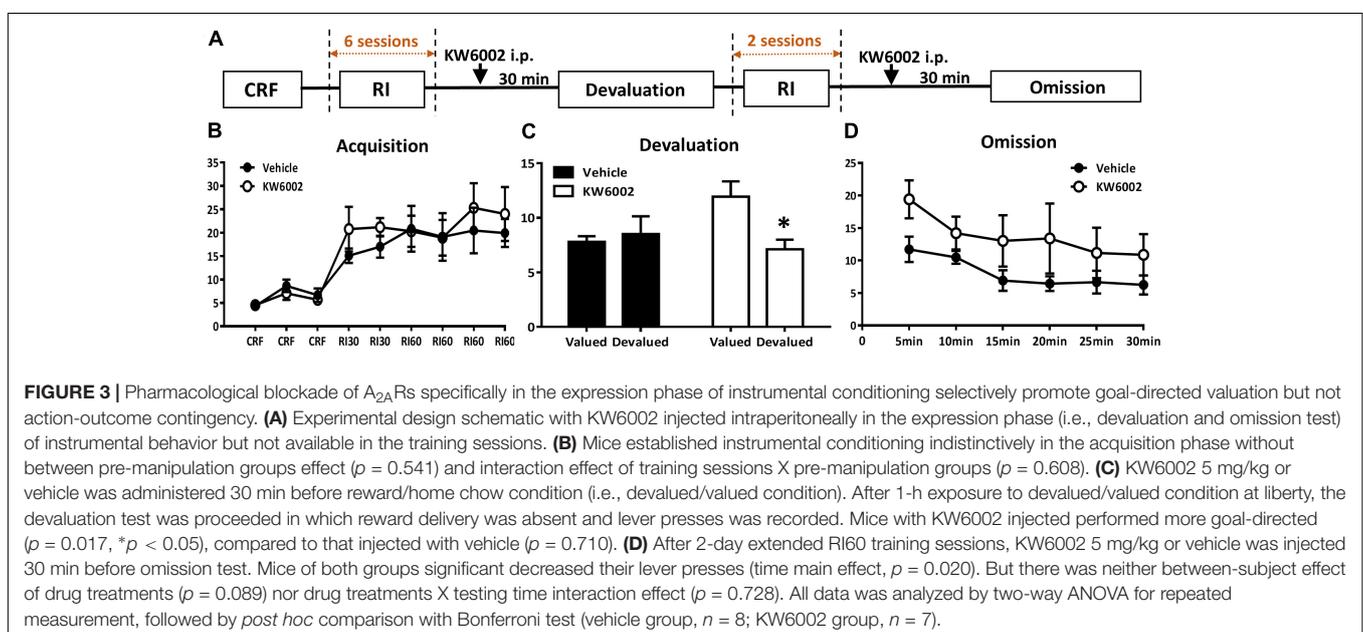
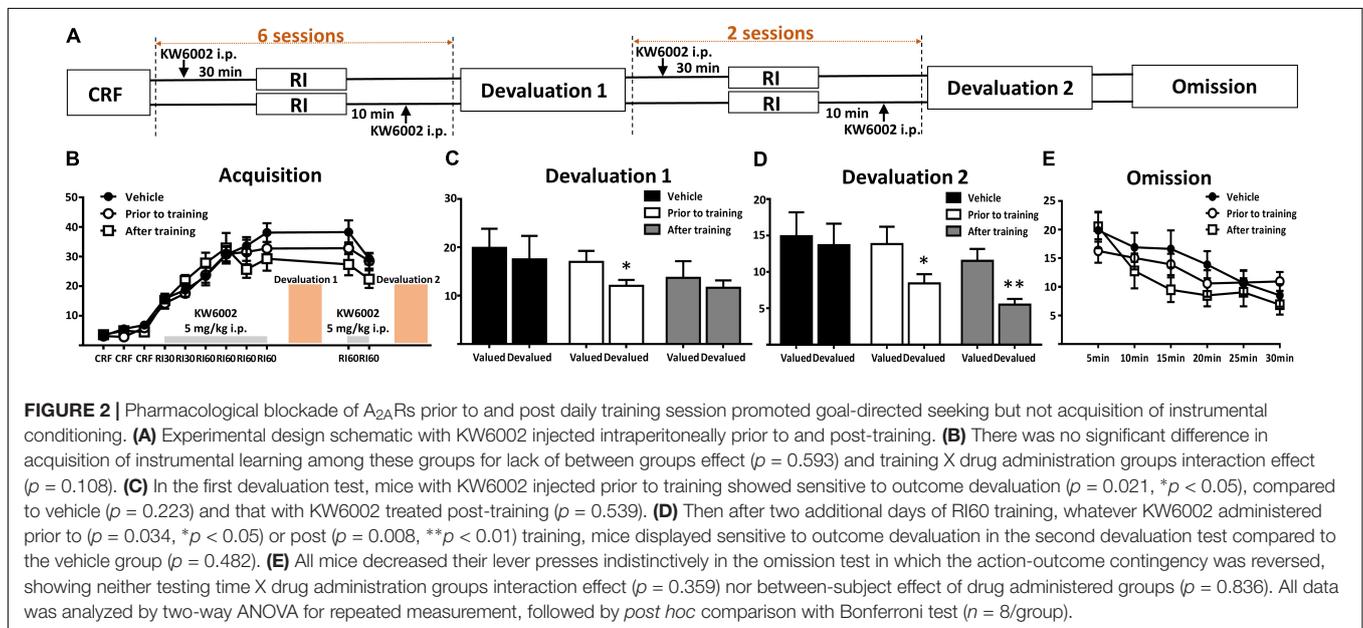
The following drugs were used in the present study: KW-6002 ((E)-1,3-diethyl-8-(3,4-dimethoxystyryl)-7-methyl-3,7-dihydro-1H-purine-2,6-dione, a selective adenosine A<sub>2A</sub>R antagonist) and DPCPX (8-cyclopentyl-1,3-dipropylxanthine, a selective adenosine A<sub>1</sub>R antagonist). KW-6002 (1 mg/kg, 5 mg/kg, Sundia, United States) was suspended in dimethyl sulfoxide (DMSO, sigma), ethoxylated castor oil (Sigma) and water with a proportion of 15%:15%:70%. DPCPX (6 mg/kg, Abcam) was dissolved in 0.9% NaCl with 5% DMSO. The control mice were treated with corresponding vehicles. All the solutions were prepared immediately before administration. The administered doses of KW-6002 and DPCPX referred to previous researches (Chen et al., 2001; Prediger et al., 2004; Nguyen et al., 2014). Drugs were injected intraperitoneally (i.p.) routinely in a volume

of 0.1 ml/10 g of body weight. The specific drug administration time course depended on experimental designs: prior to (30 min before) and post (10 min after) everyday RI training for learning and consolidation periods of instrumental learning, respectively (Figure 2A), while treated 30 min before devaluation test/omission test, but not available in the RI training sessions for expression of instrumental behavior (Figure 3A).

## DPCPX Concentration Detection

Considering the critical role of the striatum in control of instrumental behavior, we measured the concentration of DPCPX in the striatum of mice after intraperitoneal injection to verify the effective concentration of DPCPX. 30 min after DPCPX

(6 mg/kg, i.p.) administration, the striata of mice were collected and homogenized. 0.1 ml of collected homogenate was added to a 1.5 ml centrifuge tube and followed by the addition of 0.01 ml methanol and 0.3 ml of acetonitrile. The tubes were vortex mixed for 0.5 min. After centrifugation at 13,000 rpm for 10 min, 100  $\mu$ l of supernatant was transferred to an auto-sampler vial. Next, 2  $\mu$ l of the mixture was injected into the LC-MS/MS system for analysis. DPCPX concentrations were determined by ultrahigh performance liquid chromatography with mass spectrometry method (UHPLC-MS/MS). UHPLC-MS/MS analyses were performed by an Agilent UHPLC unit (Agilent Corporation, MA, United States) with a ZORBAX Eclipse Plus C18 column (1.8  $\mu$ m, 2.1  $\times$  50 mm, I.D. Agilent



Corporation, MA, United States) thermostated at 25°C. The mobile phase was composed of 0.1% formic acid (A) and acetonitrile (B) with gradient as follows: 0.0 min at 50% B, 0.0–2.0 min linear increase to 98% B, and 2.0–3.5 min at 50% B and the flow rate was 0.4 ml/min. The total run time was 3.5 min. The electrospray interface was maintained at 500°C. Nitrogen nebulization was performed with a nitrogen flow of 800 l/h. Argon was used as the collision gas. DPCPX was detected in multiple reaction monitoring (MRM) scan mode with positive ion detection. The precursor-product ion pairs used for the MRM detection were  $m/z$  305.4 → 178.1 for DCPCX.

## Quantitative PCR of A<sub>1</sub>R mRNA

Striatal tissues from A<sub>1</sub>R KO mice and their WT littermates were analyzed by the quantitative real-time polymerase chain reaction (qPCR) procedure as we have described previously (Zhang et al., 2015) using the following forward and reverse primers for A<sub>1</sub>R mRNA: primers: forward, 5'-CATCCTGGCTC TGCTTGCTATT-3'; reverse and 5'-TTGGCTATCCAGGCTTGTTCC-3'.

## Statistical Analysis

All data presented as mean ± SEM and were processed with SPSS 17.0. Two-way ANOVA for repeated measurements was used with training/testing sessions as within-subject effect and different drug administrations/genotypes as between-subject effect, followed by *post hoc* comparison by Bonferroni test, and with  $p < 0.05$  as statistical significance.

## RESULTS

### Pharmacological Blockade of A<sub>2A</sub>Rs Promoted Goal-Directed Valuation

To perform flexible, goal-directed actions, animals must acquire the ability to encode both the *contingency* between a specific action and its outcome, and the current *value* of the outcome during instrumental conditioning (Balleine and Dickinson, 1998). We administered KW6002 (i.p. at 1 mg/kg or 5 mg/kg or vehicle) 5 min prior to everyday RI training session which was critical for establishment of habitual action (**Figure 1B**) to investigate the modulatory effect of A<sub>2A</sub>R blockade on the acquisition of instrumental behaviors. To better identify goal-directed behavioral pattern, we have also included another group of mice that were trained in parallel with RR paradigm which led to goal-directed behavior as control (**Figure 1B**). All mice gradually increased their lever presses and reached a platform eventually, indicating the successful training paradigm (**Figure 1C**). Mice treated with KW6002 at 5 mg/kg significantly elevated lever presses rate (interaction effect of training sessions X drug administration groups:  $F_{5,140} = 2.659$ ,  $p = 0.006$ ; between-subject effect of drug administration groups:  $F_{3,28} = 3.740$ ,  $p = 0.022$ ): the statistical significance was observed between the RI + KW6002 5 mg/kg and the RR + Vehicle groups (Bonferroni *post hoc* test,  $p = 0.035$ ) but absent in any other comparison pairs including RI+KW6002 5 mg/kg

versus RI + Vehicle groups (*post hoc* by Bonferroni test,  $p = 0.116$ ).

The outcome devaluation procedure was used to demonstrate the importance of the evaluative components of goal-directed actions by A<sub>2A</sub>R blockade. In the devaluation test, lever presses rates between the valued and devalued conditions were compared (**Figure 1D**). Mice in the RI + Vehicle training group did not decrease lever presses in the devalued condition, showing no devaluation effect and indicating a habitual behavior ( $F_{1,7} = 0.105$ ,  $p = 0.755$ ), while the RR + Vehicle training group significantly decreased their lever presses ( $F_{1,8} = 20.865$ ,  $p = 0.002$ ), demonstrating goal-directed behavior. Notably, KW6002 at 1 mg/kg tended to decrease lever pressing rate in devalued condition compared to valued condition ( $F_{1,6} = 2.867$ ,  $p = 0.141$ ), whereas KW6002 at 5 mg/kg group showed markedly sensitive to outcome devaluation with decreased level pressing rate ( $F_{1,7} = 7.418$ ,  $p = 0.030$ ). Thus, pharmacological blockade of A<sub>2A</sub>R promoted goal-directed valuation. Whether the A<sub>2A</sub>R antagonist influence the acquisition of the instrumental learning need further clarification since the increased lever presses rate by KW6002 in the acquisition phase might be attributed to the improvement in instrumental learning or enhanced general motor activity effect of the A<sub>2A</sub>R antagonist given the drug administration immediately (~5 min) prior to behavioral training. Additional studies with the A<sub>2A</sub>R antagonist administration 30 min prior to or post training might better dissociate the learning from motor effect of A<sub>2A</sub>R antagonist.

### Pharmacological Blockade of A<sub>2A</sub>R at the Coding, Consolidation and Expression Phases of Instrumental Behavior Exerted Its Enhanced Effect on Goal-Directed Valuation but Not on Action-Outcome Contingency

To further determine the modulatory effect of A<sub>2A</sub>R on the distinct processes of instrumental behavior (i.e., learning/coding, consolidation and expression phases), we administered KW6002 at specific time course of instrumental learning processes. Based on our previous study showing the effective biological (i.e., motor) effect of KW6002 5 mg/kg maintained for 150–170 min (Shen H.Y. et al., 2008; Yu et al., 2008), we selected the specific three time points for KW6002/vehicle administration (**Figures 2A, 3A**): (a) prior to training (30 min before RI training) or (b) post training (10 min after RI training) or (c) prior to behavioral testing (30 min before devaluation/omission test but not available in the RI training sessions) to determine the modulatory effects of KW6002 on coding and consolidation phases as well as the expression of instrumental behavior, respectively.

**Figure 2B** shows that KW6002 treatment either at the prior to-training phase or post-training phase did not affect the performance of mice during the RI sessions (main effect between drug administration groups,  $F_{2,21} = 0.536$ ,  $p = 0.593$  and training sessions X drug administration groups interaction

effect,  $F_{14,147} = 2.480$ ,  $p = 0.108$ ). In the first devaluation test (**Figure 2C**), mice with vehicle injection formed a stable habitual behavior ( $F_{1,7} = 1.787$ ,  $p = 0.223$ ) as expected. Importantly, mice injected with KW6002 prior to everyday RI training session, which is the coding period, decreased their lever presses rate remarkably in the devalued condition ( $F_{1,7} = 8.779$ ,  $p = 0.021$ ), indicating blockade of A<sub>2A</sub>R enhanced goal-directed coding. However, since KW6002 post-training group did show some trend in decreasing lever pressing rate in the devaluation test, albeit not reaching statistical significance ( $F_{1,7} = 0.417$ ,  $p = 0.539$ ), we further explore the goal-directness promoting effect by KW6002 in the consolidation phase, by proceeding a 2-day complementary RI60 training sessions after the first devaluation test. Then we performed second devaluation test as illustrated in **Figure 2A**. After 2 additional days of RI training, both prior to-training and post-training groups significantly reduced lever presses in the devalued condition (prior to-training group,  $F_{1,7} = 6.931$ ,  $p = 0.034$ ; post-training group,  $F_{1,7} = 13.413$ ,  $p = 0.008$ ), i.e., goal-directed behavior, while control group (i.e., injected with vehicle) showed the characteristics of habitual behavior ( $F_{1,7} = 0.552$ ,  $p = 0.482$ ) (**Figure 2D**). Thus, KW6002 treatment in the consolidation phase of instrumental behavior promoted goal-directed behavior as well. Lastly, we performed the omission test during which the established lever press-reward association was reversed, so reward delivery depended on withholding the lever press action. As illustrated in **Figure 2E**, all mice decreased lever presses rate indistinctively in the omission test. Neither interaction effect of testing time X drug administration groups ( $F_{10,105} = 1.124$ ,  $p = 0.359$ ) nor main effect between drug administration groups ( $F_{2,21} = 0.997$ ,  $p = 0.836$ ) were detected. Thus, blockade of A<sub>2A</sub>Rs at the coding or consolidation phases of instrumental behavior enhanced goal-directed valuation but did not affect action-outcome association.

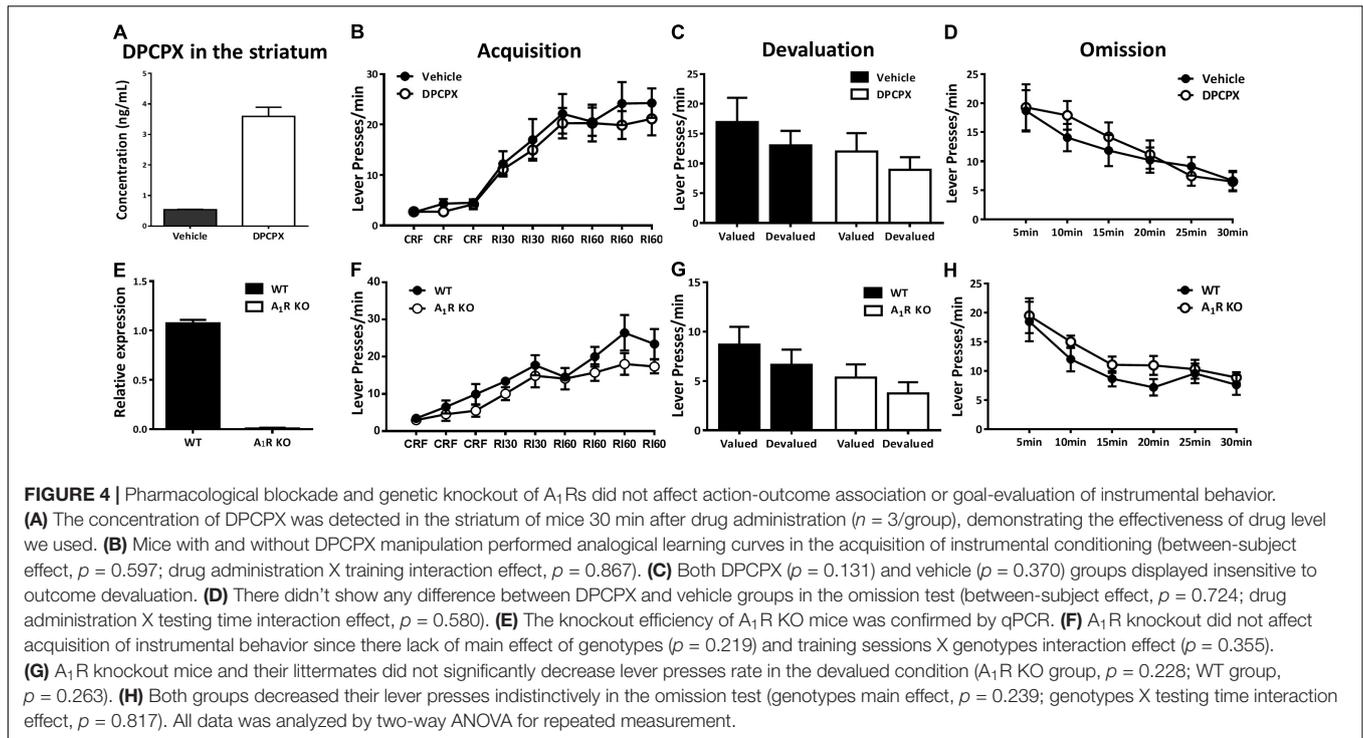
We then sought to investigate whether A<sub>2A</sub>R exerted its effect by acting on expression phase of instrumental behavior. In this experiment, KW6002 was administered 30 min before behavioral tests (devaluation and omission tests), but unavailable in all of the RI training sessions (**Figure 3A**). As expected, both pre-manipulation groups gradually increased lever presses rate and reached the platform and didn't show any difference between each other (between groups effect,  $F_{1,13} = 0.395$ ,  $p = 0.541$ ; interaction effect of training sessions X pre-manipulation groups,  $F_{5,65} = 0.554$ ,  $p = 0.608$ ) (**Figure 3B**). As **Figure 3C** shows, mice with KW6002 treatment at the expression phase displayed markedly sensitivity to outcome devaluation ( $F_{1,6} = 10.857$ ,  $p = 0.017$ ) compared with the controls ( $F_{1,7} = 0.150$ ,  $p = 0.710$ ) in the devaluation test. Thus, blockade of A<sub>2A</sub>R facilitated expression of goal-directed behavior. In the omission test (**Figure 3D**), both groups decreased their lever presses gradually over testing time (testing time main effect:  $F_{5,65} = 4.226$ ,  $p = 0.020$ ), indicating the timing effectiveness of the omission test. But the tendencies of lever-press decrease rate for the two groups were parallel as indicated by the absent of the drug treatments X testing time interaction effect ( $F_{5,65} = 0.365$ ,  $p = 0.728$ ), though mice injected with KW6002 apparently

pressed more than that of the vehicle-treated mice (between-subject effect of drug treatments,  $F_{1,13} = 3.369$ ,  $p = 0.089$ ). The increased lever presses rate by KW6002 in the omission test might attribute to general motor but not learning effect of A<sub>2A</sub>R antagonist, for drug administration was 30 min before the test. Therefore, the action-outcome contingency may not be affected by A<sub>2A</sub>R antagonist.

## Pharmacological Blockade and Genetic Knockout of A<sub>1</sub>Rs Did Not Affect Acquisition or Goal-Evaluation of Instrumental Behavior

Adenosine acts on facilitating A<sub>2A</sub>R and inhibitory A<sub>1</sub>R to integrate dopamine, glutamate, and BDNF signaling to modulate synaptic plasticity. We next investigated the possible involvement of A<sub>1</sub>Rs in the modulation of instrumental behavior. To ensure the effective DPCPX drug concentration in the striatum after our A<sub>1</sub>R pharmacological treatment paradigm, we determined the pharmacokinetic characteristic of DPCPX (**Figure 4A**) and showed the effective concentration of DPCPX in accordance with its biological effect as described previously (Baumgold et al., 1992). The A<sub>1</sub>R antagonist DPCPX (6 mg/kg) did not affect lever pressing performance during instrumental training sessions (**Figure 4B**, main effect between drug administration groups,  $F_{1,14} = 0.293$ ,  $p = 0.597$ ; interaction effect of drug administration groups X training sessions,  $F_{5,70} = 0.371$ ,  $p = 0.867$ ). The devaluation test proceeded in drug-free condition (**Figure 4C**) revealed that mice with or without DPCPX treatment responded insensitively to satiety devaluation (DPCPX group,  $F_{1,7} = 2.922$ ,  $p = 0.131$ ; vehicle group,  $F_{1,7} = 0.916$ ,  $p = 0.370$ ). In addition, both groups of mice reduced lever presses indistinguishably in the omission test (**Figure 4D**, main effect between drug administration groups,  $F_{1,14} = 0.129$ ,  $p = 0.724$ ; interaction effect of drug administration groups X testing time,  $F_{5,70} = 0.610$ ,  $p = 0.580$ ).

To further confirm this finding by pharmacological blockade of A<sub>1</sub>Rs, we determined the effect of genetic knockout of the A<sub>1</sub>R on acquisition and goal-evaluation using A<sub>1</sub>R knockout mice and their wild-type littermates. The nearly complete deletion of A<sub>1</sub>Rs was verified by qPCR (**Figure 4E**). All mice, regardless of genotypes, increased their rate of lever pressing during the training sessions (**Figure 4F**) with no significant difference between genotypes ( $F_{1,13} = 1.669$ ,  $p = 0.219$ ) or interaction between training sessions and genotypes ( $F_{5,65} = 1.105$ ,  $p = 0.355$ ). During the devaluation test (**Figure 4G**), both A<sub>1</sub>R KO and WT mice similarly showed insensitive to outcome devaluation (A<sub>1</sub>R KO group,  $F_{1,6} = 1.802$ ,  $p = 0.228$ ; WT group,  $F_{1,7} = 1.483$ ,  $p = 0.263$ ), indicating that their responding was habitual. The omission test (**Figure 4H**) further confirmed the results of pharmacological blockade of A<sub>1</sub>R by genetic knockout approach: there was neither main (genotypes) effect ( $F_{1,13} = 1.521$ ,  $p = 0.239$ ) nor the interaction of genotypes X testing time ( $F_{5,65} = 0.260$ ,  $p = 0.817$ ). This finding suggested that A<sub>1</sub>R exerted limited effect on the control of instrumental behavior.



## DISCUSSION

### A<sub>2A</sub>R Antagonist Modulate Animals' Sensitivity to Goal-Directed Valuation Without Modifying Action-Outcome Contingency

Action-outcome contingency and goal-directed valuation are two cognitive components involved in instrumental conditioning (Balleine and Dickinson, 1998). Action-outcome contingency is determined by the causal relationship between the particular actions and outcomes, while goal-directed valuation depends on the anticipation or desire for the outcome (Yin and Knowlton, 2006). Both components were acquired in the training sessions of instrumental behavior. Thus, outcome devaluation procedure was specialized to probe the importance of the evaluative component of goal-directed actions. We found that pharmacological blockade of A<sub>2A</sub>R critically promoted animals' sensitivity to outcome value (by the devaluation test) but did not affect action-outcome relationship (as manifested by similar performance in the training sessions and in the omission test). When administering 5 min prior to the training, KW6002 at 5 mg/kg apparently elevated the acquisition of learning curve. This enhancement is, however, potentially confounded by the enhanced general motor activity effect of the A<sub>2A</sub>R antagonist. Additional studies with the A<sub>2A</sub>R antagonist administering 30 min prior to or post-training can better dissociate the learning process from motor effect and clarify this issue. The selective modulation of animals' sensitivity to outcome devaluation by A<sub>2A</sub>R antagonist is in agreement with our recent finding that optogenetic activation of striatopallidal A<sub>2A</sub>R signaling in DMS

alters goal-valuation as evident by the devaluation test (Li et al., 2016). On the other hand, the lack of the effect of A<sub>2A</sub>R antagonist on the acquisition of instrumental behaviors collaborates with similar findings by genetic inactivation of striatal A<sub>2A</sub>Rs (Yu et al., 2009) and optogenetic activation of striatopallidal A<sub>2A</sub>R signaling (Li et al., 2016).

The mechanism underlying the selective modulation of goal-valuation by the A<sub>2A</sub>R is not clear. The previous study that overexpression of the D<sub>2</sub>R in the striatopallidal pathway is associated with a shift in behavioral control from habitual action to goal-directed responding but did not affect acquisition phase of instrumental learning (Kwak et al., 2014). Also, loss of striatal endocannabinoid-mediated long-term depression selectively in DLS striatopallidal neurons prevent the transition from goal-directed seeking to habitual responding behavior but did not interfere lever-press performance in the acquisition phase (Gremel et al., 2016). Given the documented antagonistic interaction of the A<sub>2A</sub>R-D<sub>2</sub>R and the A<sub>2A</sub>R-CB<sub>1</sub>R in the striatum by possibly the A<sub>2A</sub>R-D<sub>2</sub>R heterodimers (He et al., 2016) and A<sub>2A</sub>R-CB<sub>1</sub>R heterodimers (Moreno et al., 2017), these findings suggest that A<sub>2A</sub>R may selectively influence coding of the current value of the outcome (but not the contingency association) by the A<sub>2A</sub>R interaction with the D<sub>2</sub>R and CB<sub>1</sub>R functions in the striatum.

Moreover, this selective control of animals' sensitivity to reward valuation by A<sub>2A</sub>R might be related to a motivation factor, as A<sub>2A</sub>R (Mingote et al., 2008; Nam et al., 2013a) and D<sub>2</sub>R (Trifilieff et al., 2013) activities in the striatum contribute to motivational control of behaviors. Lastly, since the A<sub>2A</sub>R are predominantly expressed in the striatopallidal neurons, the A<sub>2A</sub>R control of goal-directed valuation is further

supported by the finding from the striatal circuit studies showing that as pharmacogenetic inactivation of the striatopallidal pathway enhanced motivation by energizing the initiation of goal-directed behavior (Carvalho Poyraz et al., 2016), while optogenetic stimulation of the striatopallidal pathway suppressed motivational behavior (O'Hare et al., 2016; Vicente et al., 2016).

## **A<sub>2A</sub>R Antagonist Acts at the Coding, Consolidation and Expression Phases of Instrumental Learning to Promote Goal-Directed Behavior**

Defining the specific information processing phases (i.e., learning/coding, consolidation and expression of instrumental behaviors) for A<sub>2A</sub>R antagonist control of goal-directed versus habitual behaviors is critical for our understanding of the neurotransmitter modulatory mechanisms and for the development of effective pharmacological strategy to control aberrant habit formation and drug addiction. Our demonstration of the enhanced goal-directed behavior by administration of KW6002 at the pre-training or post-training or expression phases suggests that A<sub>2A</sub>R acts at the coding, consolidation and expression phases of instrumental learning to promote animals' sensitivity to goal-directed valuation. It should be noted that the influence of the pre-training treatment paradigm on the goal-directed behavior might be partly attributed to its effect on the consolidation phase due to the relatively long-lasting effect (>2 h) of the A<sub>2A</sub>R antagonist KW6002. The similar control of instrumental behaviors by multiple treatment paradigms of KW6002 indicate that A<sub>2A</sub>R control of instrumental behaviors is largely independent of the confounding motor activity.

Various neurotransmitter systems have been implicated in control of the distinct phases of instrumental conditioning. For example, NMDA receptor signaling preferentially affected the coding (by administering NMDA antagonist at the pre-training phase) but not the expression (by administering NMDA antagonist at the post-training phase) of the instrumental conditioning (Yin et al., 2005). Furthermore, virus-induced overexpression of D<sub>2</sub>R (Trifilieff et al., 2013) and 5-HT<sub>6</sub> receptor (Eskenazi and Neumaier, 2011; Eskenazi et al., 2015) preferentially affect the coding course of operant conditioning. Additionally, optogenetic activation of endocannabinoid signaling in the training session and pharmacogenetic suppression of endocannabinoid signaling in the devaluation test gated habit formation (Gremel et al., 2016), indicating that endocannabinoid modulated instrumental learning in both coding and expression sessions, consistent with the CB<sub>1</sub>R knockout study (Hilario et al., 2007). Thus, the A<sub>2A</sub>R may interact with multiple neurotransmitter systems in the cortico-striatal projection pathways to integrate/modulate glutamate, dopamine and endocannabinoid signaling for instrumental behavioral control at multiple phases of information processing. Furthermore, cognitive control and working memory processes are important for the efficient control of goal-directed behavior (Buschman and Miller, 2014). We and others have documented that the A<sub>2A</sub>R antagonists or focal A<sub>2A</sub>R knockdown in the DMS significantly enhance working memory (Wei et al., 2014;

Kaster et al., 2015; Li et al., 2018). Thus, it is possible that when KW6002 is administered prior to the training phase, the A<sub>2A</sub>R antagonist may enhance goal-directed behavior by improving working memory. On the other hand, other mechanisms (such as "off-line" processing during sleep) may contribute to the A<sub>2A</sub>R antagonist-mediated enhancement of goal-directed behavior when A<sub>2A</sub>R antagonists are administered after the training or during the expression/retrieval phase.

## **Pharmacological Blockade and Genetic Knockout of A<sub>1</sub>Rs Did Not Affect Acquisition or Goal-Evaluation of Instrumental Behavior**

Adenosine signaling acts at the facilitating A<sub>2A</sub>R and inhibitory A<sub>1</sub>R to exert its homeostatic control of brain function. However, very limited information is available regarding the A<sub>1</sub>R control of cognition, particularly instrumental behaviors. With its relatively high expression in the cerebral cortex, hippocampus and striatum (Reppert et al., 1991; Dixon et al., 1996), A<sub>1</sub>R activation has a profound inhibitory control of excitatory transmission by presynaptic and post-synaptic mechanisms (Dunwiddie and Masino, 2001; Ribeiro et al., 2002). Striatal A<sub>1</sub>Rs can preferentially interact with the striatal D<sub>1</sub>Rs via possible A<sub>1</sub>R-D<sub>1</sub>R heterodimers in the striatonigral neurons to control striatal signaling and behavior (Gines et al., 2000). Accordingly, A<sub>1</sub>Rs modulate striatal synaptic plasticity, and prevent scopolamine- and morphine-induced impairment in working memory (Hooper et al., 1996; Lu et al., 2010). However, in the fix-interval and fix-ratio operant training paradigms, A<sub>1</sub>R antagonist failed to increase lever pressing rate, but decreased fix ratio 20 (FR20, every 20 lever presses resulted in one reward) responding at higher doses (Randall et al., 2011). Operant performance alone was insufficient to define instrumental learning modes as goal-directed or habitual actions without devaluation and omission test (Yin and Knowlton, 2006). Thus, the role of the A<sub>1</sub>R in goal-directed versus habitual behaviors is still unknown. Our study demonstrated that pharmacological blockade or global knockout of A<sub>1</sub>R did not affect the acquisition of instrumental learning or sensitivity to reward value or reversal of action-outcome relationship. This finding is in agreement with a recent study that DPCPX failed to reverse the effect of D<sub>2</sub>R antagonist on effort-relevant tasks but KW6002 and caffeine (a non-selective adenosine antagonist) can (Salamone et al., 2009). These findings suggest that A<sub>1</sub>R plays limited modulatory role in control of instrumental behavior and adenosine predominantly acts on A<sub>2A</sub>Rs but not A<sub>1</sub>Rs to modulate instrumental learning.

In summary, our study demonstrated that pharmacological blockade of A<sub>2A</sub>R but not A<sub>1</sub>R promote goal-directed behaviors by enhancing goal-directed valuation without affecting the action-outcome contingency and by acting at the coding, consolidation, and expression phases of goal-directed learning processes. These findings collaborates with our previous genetic and optogenetic studies, and with recent pharmacological studies of A<sub>2A</sub>R antagonists to control abnormal instrumental behavior in drug addiction paradigms (Nam et al., 2013a;

Pintsuk et al., 2016), providing pharmacological evidence for a therapeutic strategy to enhance goal-directed behaviors in neuropsychiatric disorders. The translational potential of A<sub>2A</sub>R antagonists is further enhanced by the recent demonstration of the safety profiles of the A<sub>2A</sub>R antagonist KW6002 in clinical phase III trials for motor benefit in >3500 Parkinson's disease patients (Chen et al., 2013) and by regular consumption of caffeine (a non-specific adenosine A<sub>2A</sub>R and A<sub>1</sub>R antagonist) by 50% world population.

## AUTHOR CONTRIBUTIONS

YL, YH, XZ, and J-FC designed the experiments. YL, XP, YH, YR, LH, ZW, and CH collected the data. YL, XP, YH,

YZ, and ZH analyzed the data. YL, XZ, and J-FC wrote the manuscript.

## FUNDING

This study was sponsored by the National Natural Science Foundation of China (Grant Nos. 81600983, 31771178, and 81600991), by the Start-up Fund from Wenzhou Medical University (Grant Nos. 89211010 and 89212012), the Zhejiang Provincial Special Funds (Grant No. 604161241), the Natural Science Foundation of Zhejiang Province of China (Grant Nos. LY15H090020, LQ16H090006, and LQ17H090005), and the Wenzhou Science and Technology Program (Grant Nos. 2016Y0725 and 2016Y0613).

## REFERENCES

- Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419. doi: 10.1016/S0028-3908(98)00033-1
- Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropharmacology* 35, 48–69. doi: 10.1038/npp.2009.131
- Baumgold, J., Nikodijevic, O., and Jacobson, K. A. (1992). Penetration of adenosine antagonists into mouse brain as determined by ex vivo binding. *Biochem. Pharmacol.* 43, 889–894. doi: 10.1016/0006-2952(92)90257-J
- Burguiere, E., Monteiro, P., Mallet, L., Feng, G., and Graybiel, A. M. (2015). Striatal circuits, habits, and implications for obsessive-compulsive disorder. *Curr. Opin. Neurobiol.* 30, 59–65. doi: 10.1016/j.conb.2014.08.008
- Buschman, T. J., and Miller, E. K. (2014). Goal-direction and top-down control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130471. doi: 10.1098/rstb.2013.0471
- Carvalho Poyraz, F., Holzner, E., Bailey, M. R., Meszaros, J., Kenney, L., Kheirbek, M. A., et al. (2016). Decreasing striatopallidal pathway function enhances motivation by energizing the initiation of goal-directed action. *J. Neurosci.* 36, 5988–6001. doi: 10.1523/JNEUROSCI.0444-16.2016
- Chen, J. F. (2014). Adenosine receptor control of cognition in normal and disease. *Int. Rev. Neurobiol.* 119, 257–307. doi: 10.1016/B978-0-12-801022-8.00012-X
- Chen, J. F., Eltzschig, H. K., and Fredholm, B. B. (2013). Adenosine receptors as drug targets—what are the challenges? *Nat. Rev. Drug Discov.* 12, 265–286. doi: 10.1038/nrd3955
- Chen, J.-F., Xu, K., Petzer, J. P., Staal, R., Xu, Y. H., Beilstein, M., et al. (2001). Neuroprotection by caffeine and A<sub>2A</sub> adenosine receptor inactivation in a model of Parkinson's disease. *J. Neurosci.* 21:RC143.
- de Wit, S., Barker, R. A., Dickinson, A. D., and Cools, R. (2011). Habitual versus goal-directed action control in Parkinson disease. *J. Cogn. Neurosci.* 23, 1218–1229. doi: 10.1162/jocn.2010.21514
- Dixon, A. K., Gubitza, A. K., Sirinathsinghji, D. J., Richardson, P. J., and Freeman, T. C. (1996). Tissue distribution of adenosine receptor mRNAs in the rat. *Br. J. Pharmacol.* 118, 1461–1468. doi: 10.1111/j.1476-5381.1996.tb15561.x
- Dolan, R. J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007
- Dunwiddie, T. V., and Masino, S. A. (2001). The role and regulation of adenosine in the central nervous system. *Annu. Rev. Neurosci.* 24, 31–55. doi: 10.1146/annurev.neuro.24.1.31
- Eskenazi, D., Brodsky, M., and Neumaier, J. F. (2015). Deconstructing 5-HT<sub>6</sub> receptor effects on striatal circuit function. *Neuroscience* 299, 97–106. doi: 10.1016/j.neuroscience.2015.04.046
- Eskenazi, D., and Neumaier, J. F. (2011). Increased expression of the 5-HT<sub>6</sub> receptor by viral mediated gene transfer into posterior but not anterior dorsomedial striatum interferes with acquisition of a discrete action-outcome task. *J. Psychopharmacol.* 25, 944–951. doi: 10.1177/0269881110388330
- Furlong, T. M., Supit, A. S., Corbit, L. H., Killcross, S., and Balleine, B. W. (2017). Pulling habits out of rats: adenosine 2A receptor antagonism in dorsomedial striatum rescues meth-amphetamine-induced deficits in goal-directed action. *Addict. Biol.* 22, 172–183. doi: 10.1111/adb.12316
- Gillan, C. M., Pappmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., et al. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *Am. J. Psychiatry* 168, 718–726. doi: 10.1176/appi.ajp.2011.10071062
- Gines, S., Hillion, J., Torvinen, M., Le Crom, S., Casado, V., Canela, E. I., et al. (2000). Dopamine D1 and adenosine A1 receptors form functionally interacting heteromeric complexes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8606–8611. doi: 10.1073/pnas.150241097
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 3), 15647–15654. doi: 10.1073/pnas.1014269108
- Graybiel, A. M., and Grafton, S. T. (2015). The striatum: where skills and habits meet. *Cold Spring Harb. Perspect. Biol.* 7:a021691. doi: 10.1101/cshperspect.a021691
- Gremel, C. M., Chancey, J. H., Atwood, B. K., Luo, G., Neve, R., Ramakrishnan, C., et al. (2016). Endocannabinoid modulation of orbitofrontal circuits gates habit formation. *Neuron* 90, 1312–1324. doi: 10.1016/j.neuron.2016.04.043
- Gremel, C. M., and Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat. Commun.* 4:2264. doi: 10.1038/ncomms3264
- He, Y., Li, Y., Chen, M., Pu, Z., Zhang, F., Chen, L., et al. (2016). Habit formation after random interval training is associated with increased adenosine A<sub>2A</sub> receptor and dopamine D<sub>2</sub> receptor heterodimers in the striatum. *Front. Mol. Neurosci.* 9:151. doi: 10.3389/fnmol.2016.00151
- Hilario, M. R., Clouse, E., Yin, H. H., and Costa, R. M. (2007). Endocannabinoid signaling is critical for habit formation. *Front. Integr. Neurosci.* 1:6. doi: 10.3389/neuro.07.006.2007
- Hooper, N., Fraser, C., and Stone, T. W. (1996). Effects of purine analogues on spontaneous alternation in mice. *Psychopharmacology (Berl)* 123, 250–257. doi: 10.1007/BF02246579
- Johansson, B., Halldner, L., Dunwiddie, T. V., Masino, S. A., Poelchen, W., Gimenez-Llort, L., et al. (2001). Hyperalgesia, anxiety, and decreased hypoxic neuroprotection in mice lacking the adenosine A<sub>1</sub> receptor. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9407–9412. doi: 10.1073/pnas.161292398
- Kaster, M. P., Machado, N. J., Silva, H. B., Nunes, A., Ardais, A. P., Santana, M., et al. (2015). Caffeine acts through neuronal adenosine A<sub>2A</sub> receptors to prevent mood and memory dysfunction triggered by chronic stress. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7833–7838. doi: 10.1073/pnas.1423088112

- Kim, H. F., and Hikosaka, O. (2015). Parallel basal ganglia circuits for voluntary and automatic behaviour to reach rewards. *Brain* 138, 1776–1800. doi: 10.1093/brain/awv134
- Kreitzer, A. C., and Malenka, R. C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 445, 643–647. doi: 10.1038/nature05506
- Kwak, S., Huh, N., Seo, J. S., Lee, J. E., Han, P. L., and Jung, M. W. (2014). Role of dopamine D2 receptors in optimizing choice strategy in a dynamic and uncertain environment. *Front. Behav. Neurosci.* 8:368. doi: 10.3389/fnbeh.2014.00368
- Lawrence, A. D., Sahakian, B. J., and Robbins, T. W. (1998). Cognitive functions and corticostriatal circuits: insights from Huntington's disease. *Trends Cogn. Sci.* 2, 379–388. doi: 10.1016/S1364-6613(98)01231-5
- Leung, B. K., and Balleine, B. W. (2013). The ventral striato-pallidal pathway mediates the effect of predictive learning on choice between goal-directed actions. *J. Neurosci.* 33, 13848–13860. doi: 10.1523/JNEUROSCI.1697-13.2013
- Li, Y., He, Y., Chen, M., Pu, Z., Chen, L., Li, P., et al. (2016). Optogenetic activation of adenosine A2A receptor signaling in the dorsomedial striatopallidal neurons suppresses goal-directed behavior. *Neuropsychopharmacology* 41, 1003–1013. doi: 10.1038/npp.2015.227
- Li, Z., Chen, X., Wang, T., Gao, Y., Li, F., Chen, L., et al. (2018). The corticostriatal adenosine A2A receptor controls maintenance and retrieval of spatial working memory. *Biol. Psychiatry* 83, 530–541. doi: 10.1016/j.biopsych.2017.07.017
- Lovinger, D. M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* 58, 951–961. doi: 10.1016/j.neuropharm.2010.01.008
- Lu, G., Zhou, Q. X., Kang, S., Li, Q. L., Zhao, L. C., Chen, J. D., et al. (2010). Chronic morphine treatment impaired hippocampal long-term potentiation and spatial memory via accumulation of extracellular adenosine acting on adenosine A1 receptors. *J. Neurosci.* 30, 5058–5070. doi: 10.1523/JNEUROSCI.0148-10.2010
- Mingote, S., Font, L., Farrar, A. M., Vontell, R., Worden, L. T., Stopper, C. M., et al. (2008). Nucleus accumbens adenosine A2A receptors regulate exertion of effort by acting on the ventral striatopallidal pathway. *J. Neurosci.* 28, 9037–9046. doi: 10.1523/JNEUROSCI.1525-08.2008
- Moreno, E., Chiarlone, A., Medrano, M., Puigdemívol, M., Bibic, L., Howell, L. A., et al. (2017). Singular location and signaling profile of adenosine A2A-cannabinoid CB1 receptor heteromers in the dorsal striatum. *Neuropsychopharmacology* 43, 964–977. doi: 10.1038/npp.2017.12
- Nam, H. W., Bruner, R. C., and Choi, D. S. (2013a). Adenosine signaling in striatal circuits and alcohol use disorders. *Mol. Cells* 36, 195–202. doi: 10.1007/s10059-013-0192-9
- Nam, H. W., Hinton, D. J., Kang, N. Y., Kim, T., Lee, M. R., Oliveros, A., et al. (2013b). 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. doi: 10.1523/JNEUROSCI.3094-12.2013
- Nguyen, M. D., Lee, S. T., Ross, A. E., Ryals, M., Choudhry, V. I., and Venton, B. J. (2014). Characterization of spontaneous, transient adenosine release in the caudate-putamen and prefrontal cortex. *PLoS One* 9:e87165. doi: 10.1371/journal.pone.0087165
- O'Hare, J. K., Ade, K. K., Sukharnikova, T., Van Hooser, S. D., Palmeri, M. L., Yin, H. H., et al. (2016). Pathway-specific striatal substrates for habitual behavior. *Neuron* 89, 472–479. doi: 10.1016/j.neuron.2015.12.032
- Ostlund, S. B., and Balleine, B. W. (2008). On habits and addiction: an associative analysis of compulsive drug seeking. *Drug. Discov. Today Dis. Models* 5, 235–245. doi: 10.1016/j.ddmod.2009.07.004
- Pintsuk, J., Borroto-Escuela, D. O., Pomierny, B., Wydra, K., Zaniewska, M., Filip, M., et al. (2016). Cocaine self-administration differentially affects allosteric A2A-D2 receptor-receptor interactions in the striatum. Relevance for cocaine use disorder. *Pharmacol. Biochem. Behav.* 144, 85–91. doi: 10.1016/j.pbb.2016.03.004
- Prediger, R. D., Batista, L. C., and Takahashi, R. N. (2004). Adenosine A1 receptors modulate the anxiolytic-like effect of ethanol in the elevated plus-maze in mice. *Eur. J. Pharmacol.* 499, 147–154. doi: 10.1016/j.ejphar.2004.07.106
- Randall, P. A., Nunes, E. J., Janniere, S. L., Stopper, C. M., Farrar, A. M., Sager, T. N., et al. (2011). Stimulant effects of adenosine antagonists on operant behavior: differential actions of selective A2A and A1 antagonists. *Psychopharmacology (Berl)* 216, 173–186. doi: 10.1007/s00213-011-2198-3
- Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., 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. doi: 10.1038/nrn2915
- Reppert, S. M., Weaver, D. R., Stehle, J. H., and Rivkees, S. A. (1991). Molecular cloning and characterization of a rat A1-adenosine receptor that is widely expressed in brain and spinal cord. *Mol. Endocrinol.* 5, 1037–1048. doi: 10.1210/mend-5-8-1037
- Ribeiro, J. A., Sebastiao, A. M., and De Mendonca, A. (2002). Adenosine receptors in the nervous system: pathophysiological implications. *Prog. Neurobiol.* 68, 377–392. doi: 10.1016/S0301-0082(02)00155-7
- Robbins, T. W., Gillan, C. M., Smith, D. G., De Wit, S., and Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn. Sci.* 16, 81–91. doi: 10.1016/j.tics.2011.11.009
- Rossi, M. A., Sukharnikova, T., Hayrapetyan, V. Y., Yang, L., and Yin, H. H. (2013). Operant self-stimulation of dopamine neurons in the substantia nigra. *PLoS One* 8:e65799. doi: 10.1371/journal.pone.0065799
- Rossi, M. A., and Yin, H. H. (2012). Methods for studying habitual behavior in mice. *Curr. Protoc. Neurosci.* 60, 8.29.1–8.29.9. doi: 10.1002/0471142301.ns0829s60
- Salamone, J. D., Farrar, A. M., Font, L., Patel, V., Schlar, D. E., Nunes, E. J., et al. (2009). Differential actions of adenosine A1 and A2A antagonists on the effort-related effects of dopamine D2 antagonism. *Behav. Brain Res.* 201, 216–222. doi: 10.1016/j.bbr.2009.02.021
- Shen, H. Y., Coelho, J. E., Ohtsuka, N., Canas, P. M., Day, Y. J., Huang, Q. Y., et al. (2008). A critical role of the adenosine A2A receptor in extrastriatal neurons in modulating psychomotor activity as revealed by opposite phenotypes of striatum and forebrain A2A receptor knock-outs. *J. Neurosci.* 28, 2970–2975. doi: 10.1523/JNEUROSCI.5255-07.2008
- Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851. doi: 10.1126/science.1160575
- Smith, K. S., and Graybiel, A. M. (2013a). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* 79, 361–374. doi: 10.1016/j.neuron.2013.05.038
- Smith, K. S., and Graybiel, A. M. (2013b). Using optogenetics to study habits. *Brain Res.* 1511, 102–114. doi: 10.1016/j.brainres.2013.01.008
- Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., and Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* 16, 966–973. doi: 10.1038/nn.3413
- Tebano, M. T., Martire, A., Potenza, R. L., Gro, C., Pepponi, R., Armida, M., et al. (2008). Adenosine A(2A) receptors are required for normal BDNF levels and BDNF-induced potentiation of synaptic transmission in the mouse hippocampus. *J. Neurochem.* 104, 279–286.
- Trifilieff, P., Feng, B., Urizar, E., Winiger, V., Ward, R. D., Taylor, K. M., et al. (2013). Increasing dopamine D2 receptor expression in the adult nucleus accumbens enhances motivation. *Mol. Psychiatry* 18, 1025–1033. doi: 10.1038/mp.2013.57
- Vicente, A. M., Galvao-Ferreira, P., Tecuapetla, F., and Costa, R. M. (2016). Direct and indirect dorsolateral striatum pathways reinforce different action strategies. *Curr. Biol.* 26, R267–R269. doi: 10.1016/j.cub.2016.02.036
- Wei, C. J., Augusto, E., Gomes, C. A., 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. doi: 10.1016/j.biopsych.2013.05.003
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512. doi: 10.1111/j.1460-9568.2005.04219.x

- Yu, C., Gupta, J., Chen, J. F., and Yin, H. H. (2009). Genetic deletion of A<sub>2A</sub> adenosine receptors in the striatum selectively impairs habit formation. *J. Neurosci.* 29, 15100–15103. doi: 10.1523/JNEUROSCI.4215-09.2009
- Yu, L., Shen, H. Y., Coelho, J. E., Araujo, I. M., Huang, Q. Y., Day, Y. J., et al. (2008). Adenosine A<sub>2A</sub> receptor antagonists exert motor and neuroprotective effects by distinct cellular mechanisms. *Ann. Neurol.* 63, 338–346. doi: 10.1002/ana.21313
- Zhang, S., Li, H., Li, B., Zhong, D., Gu, X., Tang, L., et al. (2015). Adenosine A<sub>1</sub> receptors selectively modulate oxygen-induced retinopathy at the hyperoxic and hypoxic phases by distinct cellular mechanisms. *Invest. Ophthalmol. Vis. Sci.* 56, 8108–8119. doi: 10.1167/iovs.15-17202

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Li, Pan, He, Ruan, Huang, Zhou, Hou, He, Wang, Zhang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.