Pathway-specific control of reward learning and its flexibility via selective dopamine receptors in the nucleus accumbens

Satoshi Yawata, Takashi Yamaguchi, Teruko Danjo, Takatoshi Hikida[1], and Shigetada Nakanishi[2]

Department of Systems Biology, Osaka Bioscience Institute, Suita, Osaka 565-0874, Japan

In the basal ganglia, inputs from the nucleus accumbens (NAc) are transmitted through both direct and indirect pathways and control reward-based learning. In the NAc, dopamine (DA) serves as a key neurotransmitter, modulating these two parallel pathways. This study explored how reward learning and its flexibility are controlled in a pathway-specific and DA receptor-dependent manner. We used two techniques (*i*) reversible neurotransmission blocking (RNB), in which transmission of the direct (D-RNB) or the indirect pathway (I-RNB) in the NAc on both sides of the hemispheres was selectively blocked by transmission-blocking tetanus toxin; and (*ii*) asymmetric RNB, in which transmission of the direct (D-aRNB) or the indirect pathway (I-aRNB) was unilaterally blocked by RNB techniques and the intact side of the NAc was infused with DA agonists or antagonists. Reward-based learning was assessed by measuring goal-directed learning ability based on visual cue tasks (VCTs) or response-direction tasks (RDTs). Learning flexibility was then tested by switching from a previously learned VCT to a new VCT or RDT. D-RNB mice and D1 receptor antagonist-treated D-aRNB mice showed severe impairments in learning acquisition but normal flexibility to switch from a previously learned strategy. In contrast, I-RNB mice and D2 receptor agonist-treated I-aRNB mice showed normal learning acquisition but severe impairments not only in the flexibility to the learning switch but also in the subsequent acquisition of learning a new strategy. D1 and D2 receptors thus play distinct but cooperative roles in reward learning and its flexibility in a pathway-specific manner.

goal-directed behavior | perseveration | transmission modulation | neurotransmission blockade | striatal neurons

The basal ganglia are the key neural substrates that control reward-based learning and its flexibility to effectively acquire rewards under changing environmental circumstances (1–3). Dysfunction of the basal ganglia leads to severe cognitive and learning impairments as exemplified in Parkinson's disease, schizophrenia, and drug addiction (4–6). In the basal ganglia circuitry, the projection neurons in the striatum and the nucleus accumbens (NAc), which is the ventral part of the striatum, are divided into two subpopulations, i.e., striatonigral neurons of the direct pathway and striatopallidal neurons of the indirect pathway (1, 7, 8). The outputs of these two parallel pathways converge at the substantia nigra pars reticulata/ventral tegmental area (VTA) and control the dynamic balance of the basal ganglia–thalamocortical circuitry (1, 9). The two types of the striatal projection neurons are morphologically indistinguishable, but the striatonigral and striatopallidal neurons selectively express D1 and D2 dopamine (DA) receptors (1, 10, 11). This difference in expression profile as well as the distinct ligand affinities of these two DA receptors is thought to be critical for modulating transmission of the pathways involved in rewarding behaviors (3, 12). However, whether and how different types of DA receptors in the two parallel pathways control reward learning and its flexibility are questions largely remaining to be answered.

In our previous study, we developed a gene-manipulating technique referred to as reversible neurotransmission blocking (RNB), in which neurotransmission in a specific neural pathway

is reversibly blocked in a doxycycline-regulated manner (13–15). In this technique, the transmission-blocking tetanus toxin is expressed in a pathway-specific and doxycycline-regulated manner, thus allowing separate and reversible blockade of neurotransmission in the direct pathway (D-RNB mice) or the indirect pathway (I-RNB mice) in vivo (15, 16). The function of the basal ganglia circuitry becomes defective only when both sides of the basal ganglia circuit are simultaneously impaired in the brain hemispheres (17, 18). We thus extended the RNB technique to an asymmetric RNB (aRNB) technique, in which one side of the basal ganglia circuit is blocked by the RNB technique, and the other intact side is treated with an agonist or antagonist specific for D1 or D2 receptors. This aRNB technique allowed us to disclose what type of DA receptors is responsible for pathway-specific modulation of rewarding behaviors. Here, we report that D1 and D2 receptors play a distinct and critical role in reward learning and its flexibility in a pathway-specific manner.

Results

Function of the Two Pathways in Visual Cue-Based Reward Learning. In the RNB mice, the expression of transmission-blocking tetanus toxin (TN) is driven by interaction of the tetracycline-responsive element (TRE) and the tetracycline-repressive transcription factor (tTA) (13–16) (Fig. 1*A*). The separate expression of tTA in either pathway is achieved by using the adeno-associated virus (AAV)–mediated gene-expression system, in which the expression of tTA is restricted in the direct and indirect pathways by the substance P promoter and the enkephalin promoter, respectively (15). Recombinant AAVs were bilaterally injected into the NAc (15), and 2–3 wk after the viral injection, we tested animal behaviors to investigate how the direct and indirect pathways were involved in reward-based learning and learning flexibility.

We first examined the learning ability of D-RNB and I-RNB mice to gain a reward in a visual cue task (VCT) (Fig. 1*B*). In this task, mice were randomly placed in one of two arms of a plus maze. The mice had to learn to make a correct left or right turn on the basis of visual cues to gain a reward placed at the terminal of the fixed arm. The control mice (wt mice) and I-RNB mice both progressively learned the correct choice by repeated training and reached more than 90% correct choices by the fifth session in the VCT (Fig. 1*C*). In contrast, D-RNB mice were impaired in correct choices throughout repeated sessions of the training. The wt mice and I-RNB mice reached the criterion for learning acquisition in 4.4 ± 0.2 and 5.4 ± 0.4 sessions, respectively (Fig. 1*D*). There was no statistical difference in acquisition criterion between these two groups. In contrast, the D-RNB mice reached the criterion in 8.0 ± 0.5 sessions, and
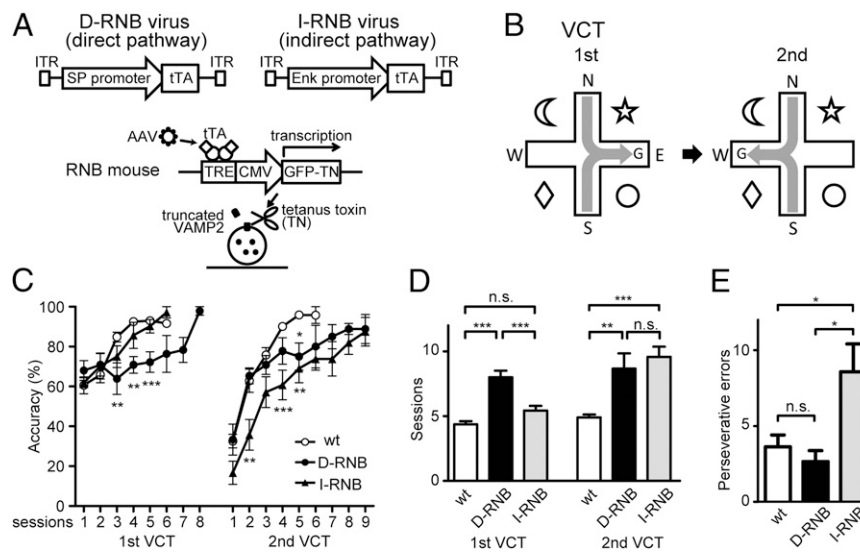
Fig. 1. Reward-based learning and its flexibility of D- and I-RNB mice in the VCT. (*A*) Schema showing preparation of D- and I-RNB mice. The D- and I-RNB viruses contained the *tTA* gene following the substance P (SP) and enkephalin (Enk) promoters, respectively. When the NAc of RNB mice was transfected with the recombinant virus, the GFP-TN fusion protein was selectively expressed in the neurons of the direct or the indirect pathway by the interaction of the virus-driven tTA with the TRE and separately blocked transmission in the respective pathway. CMV is cytomegalovirus promoter and ITR is inverted terminal repeat of the viral DNA. (*B*) Schema of learning analysis with the VCT. In the first VCT, mice were started from the north or south arm and on the basis of visual cues, they learned the east arm to receive a reward. After learning sufficiently in the first VCT, the goal (G) position was changed to the opposite, west arm in the second VCT. (*C*) Accuracy represents percentages of trials in which the mice succeeded in turning correctly to receive a reward at each session. (*D*) Number of sessions that animals required to reach the criterion in the VCT. (*E*) Number of perseverative errors on the shift of the reward position in the second task; n = 11 (wt), 6 (D-RNB), and 7 (I-RNB). Marks/columns and bars represent the mean and ± SEM, respectively. *P < 0.05, **P < 0.01, ***P < 0.001, n.s., not significant.

a significant difference in reaching the criterion was noted between the wt/I-RNB mice and D-RNB mice (*P* < 0.001).

Next we addressed how blockade of each pathway would affect the ability to learn the shift of a reward position in the VCT. In this test, animals were trained to reach the acquisition criterion in the first VCT. The reward was then placed at the end of the opposite arm so that the animal needed to make a reverse turn to receive a reward in the second task (Fig. 1*B*). In this test, the wt and D-RNB mice showed the comparable ability to learn this shift at the early sessions of training, but the learning ability of the D-RNB mice was significantly reduced at the late sessions of training (Fig. 1*C*). The I-RNB mice were impaired not only at the early sessions but also at the late sessions of training. In the second test, the wt mice achieved the learning criterion in 4.9 ± 0.2 sessions, but the D-RNB and I-RNB mice needed more training to reach it, requiring 8.7 ± 1.2 and 9.6 ± 0.8 sessions, respectively (*P* < 0.01, D-RNB vs. wt; *P* < 0.001, I-RNB vs. wt) (Fig. 1*D*).

We then addressed whether blockade of each pathway would affect perseveration after switching the VCT (Fig. 1*E*). In this analysis, perseverative errors were assessed by analyzing the trial numbers required for making a first correct turn when the reward position had been switched on the second VCT. Upon this analysis, the wt and D-RNB mice turned to the error side repeatedly at 3.6 ± 0.8 and 2.7 ± 0.7 trials, respectively; and these perseverative errors were not significantly different between these two groups. In contrast, perseverative errors by the I-RNB mice increased to 8.6 ± 1.8 trials, and this increase was significant (*P* < 0.05) compared with the trial numbers for the wt and D-RNB mice. Thus, the transmission blockade of the indirect pathway, but not that of the direct pathway, impaired learning on the switch due to perseveration.

Pathway-Dependent Function in Response-Direction Learning and Its Flexibility. The role of each pathway in reward learning was further examined by performing a response-direction task (RDT). In this task, a mouse was randomly started from two of the four

arms and had to make a 90° turn in the same direction to receive a reward (Fig. 2*A*). The wt and I-RNB mice comparably learned to make a correct turn and reached the acquisition criterion in 5.3 ± 0.5 and 5.3 ± 0.8 sessions, respectively, with no statistical difference between them (Fig. 2*B*). In contrast, the D-RNB mice showed reduced learning ability, as the number of sessions to reach the criterion significantly increased to 11.8 ± 0.9 sessions (*P* < 0.001, compared with the wt/I-RNB mice) (Fig. 2*B*). Thus, the transmission blockade of the direct pathway, as noted in the VCT test, selectively impaired acquisition of reward-based learning in the RDT test as well.

We next analyzed the ability of animals to switch to a different type of reward-based learning. When animals reached the criterion in the VCT, the task was changed to the RDT in the second test (Fig. 2*C*). Similar to the learning shift in the VCT, the D-RNB mice could learn the shift comparably as the wt mice at the early sessions of the RDT and then appeared to be partially impaired at the late sessions. It is important to note that the I-RNB mice were defective at both the early and late sessions of the learning switch. The I-RNB mice thus required a larger number of sessions (11.3 ± 2.1 sessions) to reach the criterion than did the wt mice (4.7 ± 0.6 sessions, *P* < 0.01) (Fig. 2*D*). This session number of the D-RNB mice also tended to increase (7.4 ± 0.9 sessions) although this number was not statistically significant compared with that for the wt mice (*P* = 0.14). We then analyzed perseverative errors in the VCT–RDT switching task. These errors significantly increased in the I-RNB mice (6.5 ± 1.2 trials) (*P* < 0.05), but not in the D-RNB mice (3.9 ± 0.6 trials) compared with the number for the wt mice (2.7 ± 0.6 trials) (Fig. 2*E*). Thus, blockade of the indirect pathway increased perseveration and impaired a different type of reward-based learning shift.

D1-Receptor Regulation of the Direct Pathway in Acquisition of Reward-Based Learning. To assess how DA could regulate reward-based learning and its flexibility, we generated asymmetric RNB mice, in which transmission of either the direct or the
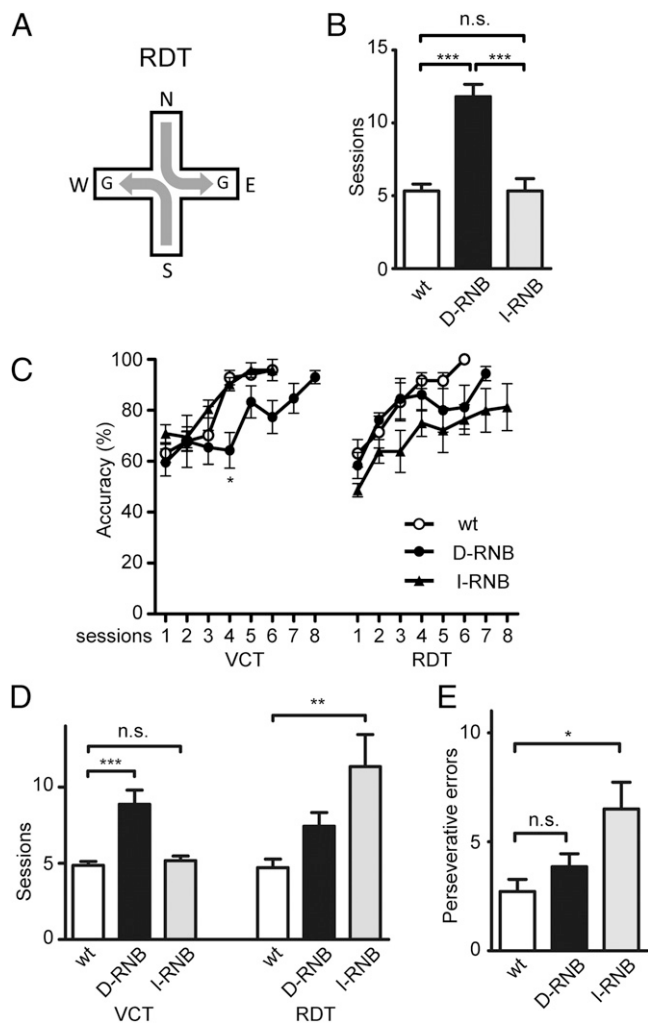
Fig. 2. Learning acquisition in the RDT and learning flexibility on the shift from the VCT to the RDT. (*A*) Schema of learning analysis with the RDT. Mice were started from either the north or south arm and learned a fixed turning direction to receive a reward. (*B*) Number of sessions that the animals required to reach the criterion in the RDT; *n* = 9 (wt), 5 (D-RNB), and 6 (I-RNB). (*C–E*) Learning ability was tested by the VCT in the first test. After sufficiently learning in the first test, the mice were then examined for learning ability in the second test using RDT. Accuracy (*C*), session numbers to reach the criterion (*D*), and perseverative errors (*E*) were determined as described in Fig. 1; *n* = 7 (wt), 7 (D-RNB), and 6 (I-RNB). Marks/columns and bars represent the mean and ± SEM, respectively. *P < 0.05, **P < 0.01, ***P < 0.001, n.s., not significant.

indirect pathway in the NAc was unilaterally blocked by the D- or I-RNB virus injection (D-aRNB and I-aRNB mice), respectively (Fig. 3). Two to three weeks after the viral injection, a DA agonist or antagonist was infused into the intact side of the NAc through an implanted cannula (Fig. 3*A*). Animal behavior was analyzed 15–30 min after drug infusion in each session, and the location of the implanted cannula was confirmed once the behavioral analysis had been completed. SKF81297 (SKF) and SCH23390 (SCH) were used as a D1 agonist and a D1 antagonist, respectively; whereas quinpirole and eticlopride were used as a D2 agonist and a D2 antagonist, respectively (19–22).

The D1 receptor is predominantly expressed in the striatonigral neurons of the direct pathway (11, 23, 24). We first addressed whether and how D1 receptors could be involved in reward-based learning and its flexibility. In this analysis, we performed the VCT test to examine the effects of treatment of D-

aRNB mice with either saline, SCH, or SKF (Fig. 3*A*). The saline-injected D-aRNB mice showed normal learning acquisition in the first VCT test and normal learning switch in the second VCT test, as the accuracies were comparable to those of the saline-injected wt mice (Fig. 3*B*). This finding verified that unilateral blockade of transmission had no effect on the reward-based learning ability. Then, when D1 receptors were inhibited by contralateral injection of SCH into the intact NAc, these mice were severely impaired in their learning ability throughout training in the first test. Furthermore, they normally learned the shift of a reward at the early sessions but became defective at the late sessions of the second test (Fig. 3*C*). The SCH–D-aRNB mice thus exhibited a defective profile identical to that of the bilaterally blocked D-RNB mice in terms of both learning acquisition and learning switch. In contrast, stimulation of D1 receptors with SKF had no obvious effects on either learning acquisition or learning switch (Fig. 3*D*). Also, the virus-transfected wt mice never showed any defect by contralateral injection of either SCH or SKF (Fig. 3 *C* and *D*). As a result, only the SCH–D-aRNB mice showed a significant increase in the session number required to reach the learning criterion in both the first and second tests. The session numbers needed to reach the criterion in the first and second tests were 5.3 ± 0.2 and 4.5 ± 0.3 for saline–wt; 5.5 ± 0.4 and 6.2 ± 0.7 for saline–D-aRNB; and 9.4 ± 1.2 and 9.4 ± 1.3 for SCH–D-aRNB (*P* < 0.05–0.001, SCH–D-aRNB vs. other groups) (Fig. 3*E*). In addition, when the D-aRNB mice were treated with either saline, SCH, or SKF, none of these mice showed an increase in perseverative errors in response to switching of the reward position (Fig. 3*F*). The results thus indicate that the activation of D1 receptors in the direct pathway plays a key role in learning acquisition but not in learning switch.

D2-Receptor Regulation of the Indirect Pathway in Flexibility of Reward-Based Learning. The D2 receptor is predominantly expressed in the striatopallidal neurons of the indirect pathway (11, 23, 24). We next assessed the role of D2 receptors in the indirect pathway by examining the effects of a D2 agonist or antagonist on the learning ability of I-aRNB mice (Fig. 4). Infusion of saline into I-aRNB mice had no effect on either learning acquisition in the VCT of the first test or on the VCT–VCT learning switch in the second test (Fig. 4*A*). The injection of the D2 agonist quinpirole into I-aRNB mice tended to rather elevate learning acquisition in the first test and then, like the bilateral blockade in I-RNB mice, markedly impaired both the early and late sessions of learning switch in the second test (Fig. 4*B*). In contrast, the D2 antagonist eticlopride had no inhibitory effect on the learning ability in either the first or second test (Fig. 4*C*). Thus, the number of sessions to reach the criterion increased in learning switch only when the D2 receptor was activated in the I-aRNB mice: the session numbers required to reach the criterion in the first and second tests were 5.0 and 5.2 ± 0.7 for saline–wt; 5.3 ± 0.4 and 5.6 ± 0.6 for saline–I-aRNB; and 4.3 ± 0.4 and 9.0 ± 1.2 for quinpirole–I-aRNB (*P* < 0.01–0.001; quinpirole–I-aRNB vs. other groups in the second test) (Fig. 4*D*). Furthermore, the perseverative errors significantly increased in the VCT–VCT switching when quinpirole (8.0 ± 1.8 trials), but not eticlopride (3.7 ± 1.1 trials), was administered to the I-aRNB mice (*P* < 0.01, quinpirole–I-aRNB vs. saline–I-aRNB or quinpirole–wt) (Fig. 4*E*). Thus, the profile of the quinpirole–I-aRNB mice in terms of defective reward learning and its flexibility was identical to that of the bilaterally blocked I-RNB mice. These results indicate that inactivation of the D2 receptor in the indirect pathway is necessary to flexibly adapt learning switching and promote the subsequent new learning.

In this investigation, we focused on the functional role of the pathway-specific DA receptors in the NAc. Both D-aRNB and I-aRNB mice showed no alteration in locomotor activity in the plus maze regardless of treatment with DA agonists or antagonists
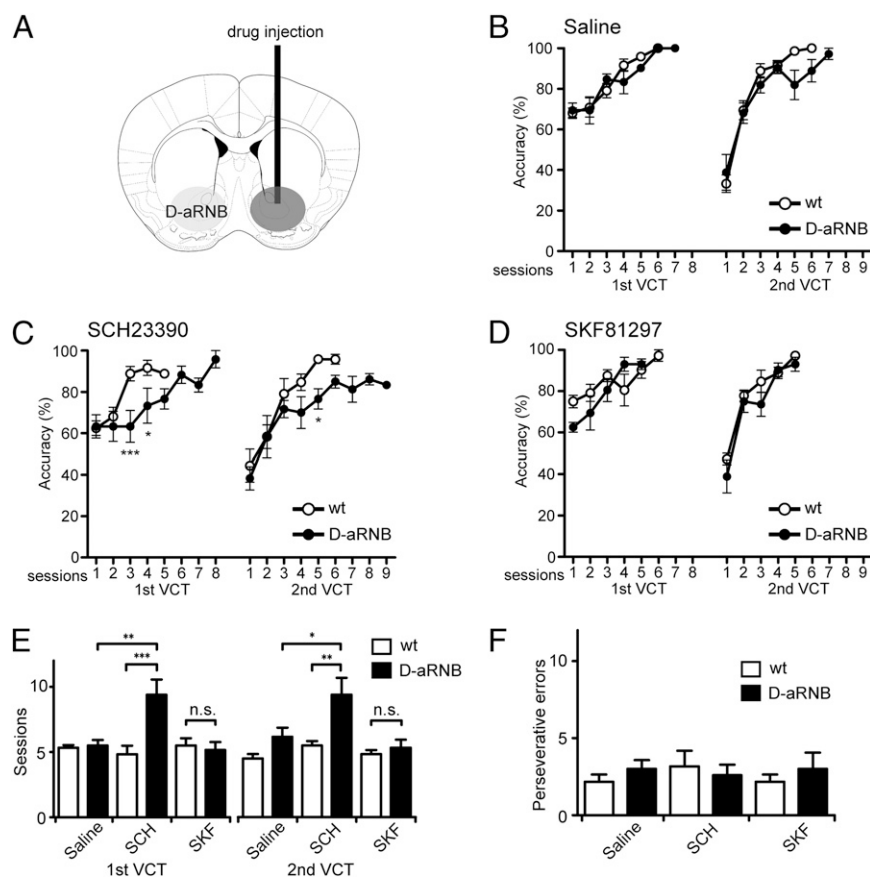
Fig. 3. Learning acquisition with D1 receptor in the direct pathway. (*A*) Schema of the aRNB technique combined with pharmacological analysis. One side of transmission of the direct pathway in the NAc was blocked by the RNB technique, and the other intact side of the NAc was infused with saline (*B*), SCH23390 (*C*), or SKF81297 (*D*). Learning ability was tested by the VCT in the first test. Then after the mice had sufficiently learned in the first test, a reward was placed at the end of the opposite arm; and the learning ability was then tested by the VCT in the second test. Accuracy (*B–D*), session numbers to reach the criterion (*E*), and perseverative errors (*F*) were determined as described in Fig. 1; $n = 5$–6. Marks/columns and bars represent the mean and $\pm$ SEM, respectively. *$P < 0.1$, **$P < 0.01$, ***$P < 0.001$, n.s., not significant.

(8.4–10.5 cm/s). Our previous study showed that when the direct or the indirect pathway was unilaterally blocked in the whole striatal region and then these mice were forced to rotate on a hemispherical container, such blockade induced abnormal ipsilateral or contralateral rotations, respectively (15). However, none of the drugs infused into the NAc elicited such abnormal turning in either the D-aRNB or I-aRNB mice. Thus, the observed impairments of learning ability in the drug-infused aRNB mice were not due to imbalance of motor movement and indeed reflected deficits in reward-based learning and its flexibility.

Discussion

This study has established a technique that allowed us to define the role of pathway-specific DA receptors in reward-based learning and its flexibility. The results explicitly indicate that learning deficits in both D1 antagonist-treated D-aRNB mice and D2 agonist-treated I-aRNB mice reflected those in bilaterally blocked D- and I-RNB mice, respectively. These results indicate that the activation of D1 receptor in the direct pathway is essential for the animals to acquire reward-based learning but is not necessary for them to flexibly switch the previously learned reward-seeking behavior. In contrast, the modulation of D2 receptors in the indirect pathway is not necessary for acquisition of reward-based learning; but their inactivation is indispensable for learning flexibility to switch from the previously learned behavior. Thus, the D1 and D2 receptors play a key role in learning acquisition and learning flexibility, respectively, in a pathway-

specific manner. The results further indicate that the indirect pathway-defective naive mice normally learned a reward-gaining strategy but that these mice, once having learned it, showed difficulty in learning the switch to a new strategy even after repeated reward presentation. Furthermore, the observed functional deficits were caused by restricted blockade of the NAc circuit, indicating that input convergence at the NAc is critical for both reward learning and its flexibility.

D1 and D2 receptors are almost exclusively expressed in the direct and indirect pathways, respectively, the former exhibiting low-affinity binding of DA and the latter, high-affinity binding (11, 23, 24). In addition, DA neurons in the VTA exhibit two different patterns of firings, a phasic firing and a tonic firing, which differentially modulate D1 and D2 receptors in the NAc (12, 17, 25). On the basis of these characteristic features of DA transmission, Frank proposed a neurocomputational model to explain "Go" and "No Go" signals in reward-based learning (26). Our study has provided explicit experimental evidence that is in good agreement with their proposal and has further extended our understanding of reward-based learning mechanisms, as depicted in Fig. 5. When naive animals encounter unexpected rewards or sensory signals predicting such rewards, DA neurons evoke a burst of phasic firings and increase synaptic concentrations of DA in the NAc (27, 28). This increase effectively activates the low-affinity D1 receptor and enhances the response of the NAc neurons to reward-related input, thereby triggering reward-directed learning (Fig. 5*A*). Thus, the defect of the direct
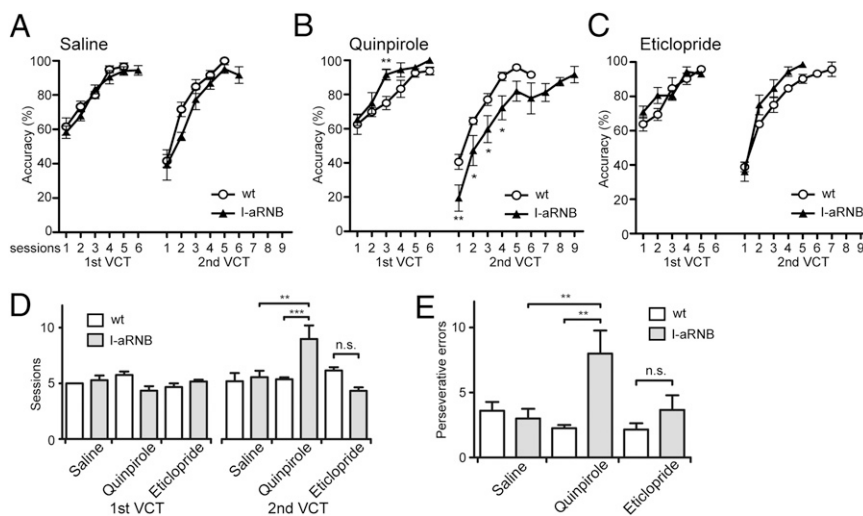
Fig. 4. Learning acquisition and flexibility with D2 receptor in the indirect pathway. One side of transmission of the indirect pathway in the NAc was blocked by the RNB technique, and the other intact side of the NAc was infused with saline (*A*), quinpirole (*B*), or eticlopride (*C*). Learning ability was tested as described in Fig. 3. Accuracy (*A–C*), session numbers to reach the criterion (*D*), and perseverative errors (*E*) were determined as described in Fig. 1; *n* = 5–8. Marks/columns and bars represent the mean and ± SEM, respectively. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, n.s., not significant.

pathway in both D-RNB mice and D1 antagonist-treated D-aRNB mice resulted in impairments of reward-based learning. The functional deficit of the D1 receptor in the direct pathway also impaired the learning ability at the late sessions of the second test. The activation of the D1 receptor is thus required for learning a new strategy when the learning strategy was switched in the second test. By contrast, absence of an expected reward suppresses tonic firings of DA neurons and lowers DA concentrations in the NAc (12). This reduction in DA relieves the D2-receptor-mediated inhibition of the indirect pathway but has no effect on the low-affinity D1 receptor in the direct pathway. The selective disinhibition of the indirect pathway then precludes the previously learned actions in response to reward omission (Fig. 5*B*). Thus, both I-RNB mice and D2-receptor-activated I-aRNB mice showed normal learning ability to initial reward presentation but were severely impaired in their flexibility of the learning switch. The important finding of this investigation is that when the learning strategy was shifted, the defective indirect pathway significantly slowed down the ability to learn a new reward-gaining strategy in both the VCT–VCT and VCT–RDT tests. When the rewarding system is changed, it is necessary to prevent recalling a previously learned strategy in addition to promoting a new reward-based learning (Fig. 5*C*). Our finding thus strongly suggests that the indirect pathway is indispensable not only for rapid suppression of perseveration toward a previously learned strategy but also to persistently preclude the reward-negative outcome (Fig. 5*C*).

The integrative processes in the NAc circuit that regulate reward learning and its flexibility have important implications for disease states where DA signaling is abnormal. In both hyperdopaminergic states, such as drug abuse and schizophrenia, and hypodopaminergic states, as seen in Parkinson's disease, the modulatory mechanisms of the direct and indirect pathways are disrupted (4–6, 29). This is reflected in the abnormal behaviors of animal models and human patients. Cocaine-sensitized animals are not impaired in acquiring reward learning but show significant perseveration if the goal is switched (30). A similar type of abnormality is observed in probabilistic reversal learning after repeated L-DOPA administration in Parkinson's patients and is proposed to be due to the lack of D2-receptor inactivation necessary for learning flexibility (29, 31). Our findings that indicate pathway-specific and receptor-dependent dopaminergic

modulation in reward learning and flexibility will provide more informative approaches for the treatment of Parkinson's disease and psychiatric disorders.

Materials and Methods

RNB Mice. All animal-handling procedures were performed according to the guidelines of Osaka Bioscience Institute. The RNB mice were generated as described previously (15), and the schema of the RNB technique is presented in Fig. 1*A*. The recombinant AAV was unilaterally or bilaterally injected into four sites of the NAc by a stereotaxic technique (15). The RNB and aRNB mice and their wt littermates were used for all experiments.

Drug Infusion in aRNB Mice. After anesthesia and retraction of the scalp, the recombinant virus was injected unilaterally into the NAc (15). Then, the
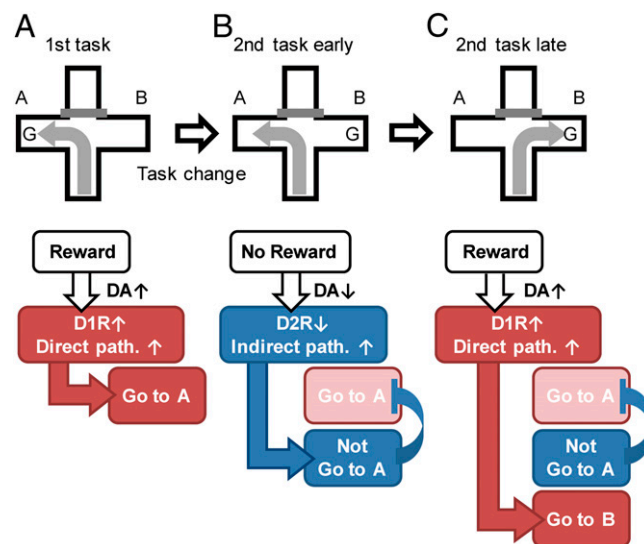

Fig. 5. Schematic model of pathway-specific, DA receptor-dependent modulation of the NAc in reward learning and flexibility. D1R is D1 receptor and D2R is D2 receptor. Note that this model holds that the D2 receptor in the indirect pathway is involved in not only the flexibility to learning switch but also the subsequent suppression of learning conflicts between a previously learned strategy and a new one.

contralateral NAc was implanted with a 5-mm guide cannula (26-gauge) possessing a dummy cannula (33-gauge) aimed toward the NAc. The guide cannula was secured in place with dental acrylic. The stereotaxic coordinates for drug infusion were 1.2 mm anterior to bregma, 1.2 mm lateral to the midline, and 3.5 mm ventral to dura according to the atlas of Franklin and Paxinos (32). Drug infusion into the NAc was made through an inner cannula (33-gauge) attached to a Hamilton syringe. The syringe was driven in a volume of 1 μL per side for 2 min by a microinfusion pump. The concentrations of infused drugs were 100 μM SCH23390, 300 μM SKF81297, 1 mM quinpirole, and 400 μM eticlopride, all purchased from Sigma. After behavioral analysis, injection sites of drugs were confirmed by direct visualization of a series of slice sections of the NAc region. When the injection site was found to be outside the NAc, these data were discarded (about 3% of drug-injected aRNB mice).

Behavioral Analysis. A four-arm cross maze was made of a clear plastic wall and a gray floor and placed 90 cm above the floor. Each arm was 25 cm long and 5 cm wide, and the center platform was 5 × 5 cm. Visual cues such as balls and shopping baskets were hung outside the maze, and one side of the room and the other side of it were surrounded by a black and a yellow curtain, respectively. The position of a mouse was detected by video camera suspended over the maze and was analyzed by use of Labview software. Behavioral analysis was started 2–3 wk after manipulation with either bilateral viral injection or unilateral viral injection together with surgery for drug infusion. Animals were food-restricted to reach approximately 80% of their original ad libitum weight by the beginning of behavioral analysis, which was started after 3 d of habituation. On each day of habituation, three pieces of chocolate were placed in the food well of each arm. A mouse was allowed to freely navigate and consume the chocolate within 15 min. During the habituation period, the mouse was handled for 10 min per day. After the habituation procedure, a possible bias to turn to a preferred arm or to a preferred direction was assessed in the absence of chocolate by the

use of a T maze, in which one arm was closed by a clear acrylic wall. Then, to avoid the possible turning bias, a reward was placed on the opposite arm as its turn bias during testing (17, 33). In the VCT, a mouse was started from either the north or south arm and had to make a 90° turn to the left or to the right on the basis of visual cues. In the RDT, a mouse had to make a 90° turn in the same direction. Each start arm was used with an equal number of trials in a pseudorandom fashion. Two sessions, each consisting 12 trials, were carried out per day. Between trials, the mouse was placed back in the holding cage. The maze arms were wiped down with a sponge moisturized with ammonium chloride solution. The intertrial interval was ~10 s. Accuracy was calculated as percentages of correct choices per session. The acquisition criterion was defined as more than 11 correct choices in two consecutive sessions. Regardless of reaching this criterion, at least five sessions were performed in each test. Perseverative errors were calculated as number of repeated incorrect choices in the beginning of the second test in which the first VCT was switched to either the second VCT or the RDT. Locomotor activity and forced rotation were measured as described previously (15). All tests of animal behaviors were conducted in a blind fashion.

Statistical Analysis. Statistical analysis was conducted by using GraphPad PRISM 5.0 (GraphPad Software) and StatView 5.0. Data were analyzed by one-way ANOVA or repeated measure ANOVA and post hoc comparisons were made by using the Bonferroni test.

1. Graybiel AM (2000) The basal ganglia. *Curr Biol* 10:R509–R511.
2. Wickens JR, Reynolds JNJ, Hyland BI (2003) Neural mechanisms of reward-related motor learning. *Curr Opin Neurobiol* 13:685–690.
3. Grace AA, Floresco SB, Goto Y, Lodge DJ (2007) Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci* 30:220–227.
4. Hyman SE, Malenka RC, Nestler EJ (2006) Neural mechanisms of addiction: The role of reward-related learning and memory. *Annu Rev Neurosci* 29:565–598.
5. Israel Z, Bergman H (2008) Pathophysiology of the basal ganglia and movement disorders: From animal models to human clinical applications. *Neurosci Biobehav Rev* 32:367–377.
6. Simpson EH, Kellendonk C, Kandel E (2010) A possible role for the striatum in the pathogenesis of the cognitive symptoms of schizophrenia. *Neuron* 65:585–596.
7. Albin RL, Young AB, Penney JB (1989) The functional anatomy of basal ganglia disorders. *Trends Neurosci* 12:366–375.
8. Alexander GE, Crutcher MD (1990) Functional architecture of basal ganglia circuits: Neural substrates of parallel processing. *Trends Neurosci* 13:266–271.
9. Deniau JM, Mailly P, Maurice N, Charpier S (2007) The pars reticulata of the substantia nigra: A window to basal ganglia output. *Prog Brain Res* 160:151–172.
10. Gerfen CR, et al. (1990) D_1 and D_2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* 250:1429–1432.
11. Surmeier DJ, Ding J, Day M, Wang Z, Shen W (2007) D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci* 30:228–235.
12. Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron* 68:815–834.
13. Yamamoto M, et al. (2003) Reversible suppression of glutamatergic neurotransmission of cerebellar granule cells in vivo by genetically manipulated expression of tetanus neurotoxin light chain. *J Neurosci* 23:6759–6767.
14. Wada N, et al. (2007) Conditioned eyeblink learning is formed and stored without cerebellar granule cell transmission. *Proc Natl Acad Sci USA* 104:16690–16695.
15. Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi S (2010) Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* 66:896–907.
16. Kimura K, Hikida T, Yawata S, Yamaguchi T, Nakanishi S (2011) Pathway-specific engagement of ephrinA5-EphA4/EphA5 system of the substantia nigra pars reticulata in cocaine-induced responses. *Proc Natl Acad Sci USA* 108:9981–9986.
17. Goto Y, Grace AA (2005) Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* 8:805–812.
18. Block AE, Dhanji H, Thompson-Tardif SF, Floresco SB (2007) Thalamic-prefrontal cortical-ventral striatal circuitry mediates dissociable components of strategy set shifting. *Cereb Cortex* 17:1625–1636.
19. Tsuruta K, et al. (1981) Evidence that LY-141865 specifically stimulates the D-2 dopamine receptor. *Nature* 292:463–465.
20. Hyttel J (1983) SCH 23390 - the first selective dopamine D-1 antagonist. *Eur J Pharmacol* 91:153–154.
21. Hall H, Köhler C, Gawell L (1985) Some in vitro receptor binding properties of [³H] eticlopride, a novel substituted benzamide, selective for dopamine-D_2 receptors in the rat brain. *Eur J Pharmacol* 111:191–199.
22. Arnt J, Bøgesø KP, Hyttel J, Meier E (1988) Relative dopamine D_1 and D_2 receptor affinity and efficacy determine whether dopamine agonists induce hyperactivity or oral stereotypy in rats. *Pharmacol Toxicol* 62:121–130.
23. Lobo MK, Karsten SL, Gray M, Geschwind DH, Yang XW (2006) FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. *Nat Neurosci* 9:443–452.
24. Heiman M, et al. (2008) A translational profiling approach for the molecular characterization of CNS cell types. *Cell* 135:738–748.
25. Shen W, Flajolet M, Greengard P, Surmeier DJ (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321:848–851.
26. Frank MJ (2011) Computational models of motivated action selection in corticostriatal circuits. *Curr Opin Neurobiol* 21:381–386.
27. Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72:1024–1027.
28. Roitman MF, Wheeler RA, Wightman RM, Carelli RM (2008) Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat Neurosci* 11:1376–1377.
29. Maia TV, Frank MJ (2011) From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14:154–162.
30. Goto Y, Grace AA (2005) Dopamine-dependent interactions between limbic and prefrontal cortical plasticity in the nucleus accumbens: Disruption by cocaine sensitization. *Neuron* 47:255–266.
31. Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007) Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* 318:1309–1312.
32. Franklin KBJ, Paxinos G (2008) *The Mouse Brain in Stereotaxic Coordinates* (Academic, San Diego).
33. Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP (2002) Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav Neurosci* 116:105–115.

NEUROSCIENCE