Behavioral/Systems/Cognitive

Acquisition and Performance of Goal-Directed Instrumental Actions Depends on ERK Signaling in Distinct Regions of Dorsal Striatum in Rats

Michael W. Shiflett,[1,2] **Robert A. Brown,**[1] **and Bernard W. Balleine**[1,3]

[1]Department of Psychology and the Brain Research Institute, University of California, Los Angeles, Los Angeles, California 90095-1563, [2]Department of Psychology, Rutgers University, Newark, New Jersey 07102, and [3]Brain and Mind Research Institute, University of Sydney, Sydney, New South Wales 2050, Australia

The performance of goal-directed actions relies on an animal's previous knowledge of the outcomes or consequences that result from its actions. Additionally, a sensorimotor learning process linking environmental stimuli with actions influences instrumental performance by selecting actions for additional evaluation. These distinct decision-making processes in rodents depend on separate subregions of the dorsal striatum. Whereas the posterior dorsomedial striatum (pDMS) is required for the encoding of actions with their outcomes or consequences, the dorsolateral striatum (DLS) mediates action selection based on sensorimotor learning. However, the molecular mechanisms within these brain regions that support learning and performance of goal-directed behavior are not known. Here we show that activation of extracellular signal-regulated kinase (ERK) in the dorsal striatum has a critical role in learning and performance of instrumental goal-directed behavior in rodents. We observed an increase in p42 ERK (ERK2) activation in both the pDMS and DLS during both the acquisition and performance of recently acquired instrumental goal-directed actions. Furthermore, disruption of ERK activation in the pDMS prevented both the acquisition of action–outcome associations, as well as the performance of goal-directed actions guided by previously acquired associations, whereas disruption of ERK activation in the DLS disrupted instrumental performance but left instrumental action–outcome learning intact. These results provide evidence of a critical, region-specific role for ERK signaling in the dorsal striatum during the acquisition of instrumental learning and suggest that processes sensitive to ERK signaling within these striatal subregions interact to control instrumental performance after initial acquisition.

Introduction

Evidence from instrumental conditioning in rats suggests that choice between different courses of action depends on integrating causal knowledge of the relationship between actions and their consequences with the current incentive value of those consequences (Dickinson and Balleine, 1994; Balleine and Dickinson, 1998a). Both posttraining changes in the incentive value of the instrumental outcome and treatments that degrade the instrumental action–outcome contingency attenuate the rate of performance of an action and modify the rats' choice between actions (Balleine and Dickinson, 1998a,b). Recently, it has become clear that an additional sensorimotor learning process linking environmental stimuli with actions influences performance by selecting actions for additional evaluation (Balleine and Ostlund, 2007; Ostlund and Balleine, 2007). When actions are overtrained, these stimulus–response associations can elicit actions directly and independently of their consequences, ren-

dering performance inflexible or habitual (Dickinson, 1994; Dayan and Balleine, 2002).

Recent experiments have revealed that different subregions of the dorsal striatum mediate these distinct decision-making processes in rodents. Lesions within a posterior region of dorsomedial striatum (pDMS) abolish goal-directed learning and render choice performance insensitive to contingency degradation and outcome devaluation treatments, i.e., choice becomes rigid and habitual (Yin et al., 2005b). A parallel corticostriatal circuit involving the dorsolateral striatum (DLS) in rodents mediates action selection based on sensorimotor learning. Whereas overtraining causes performance to become insensitive to outcome devaluation and contingency degradation, lesions and temporary inactivation of DLS reverse this effect, rendering performance again sensitive to these treatments (Yin et al., 2004). This evidence supports the general claim that distinct corticostriatal networks control different aspects of the decision process (Daw et al., 2005; Yin et al., 2006).

Nevertheless, the molecular mechanisms that underlie instrumental learning and performance are not well understood. Disruption of dopamine and glutamate signaling within the striatum interferes with instrumental learning and performance and prevents long-term potentiation (LTP) of corticostriatal synapses, a process thought to be necessary for instrumental learning (Reynolds et al., 2001; Andrzejewski et al., 2004; Faure et al., 2005; Yin

et al., 2005a; Dang et al., 2006; Di Filippo et al., 2009). Corticostriatal LTP requires activation of extracellular signal-regulated kinase (ERK) a member of the mitogen-activated protein (MAP) kinase pathway (Sgambato et al., 1998; Mazzucchelli et al., 2002). Overexpression of the p42 isoform of ERK (ERK2) in striatum enhances corticostriatal LTP, memory retention in active and passive avoidance tasks, and expression of drug conditioned place preference (CPP), which suggests that ERK2 activation in the striatum may play a key role in instrumental learning and performance (Mazzucchelli et al., 2002; Ferguson et al., 2006). In the present study, we examined ERK2 activation in the dorsal striatum during both the acquisition of specific action–outcome associations and during their utilization in choice performance. We hypothesized that ERK2 signaling would be differentially activated in the pDMS and DLS after different amounts of instrumental training, and, furthermore, that disruption of ERK activation in these structures would have different effects on instrumental learning and performance, consistent with the involvement of these regions in action–outcome and sensorimotor learning.

Materials and Methods

Subjects and apparatus
A total of 87 adult male Long–Evans rats (Harlan) were used in this study. Rats arrived weighing 250–275 g and were housed individually in Plexiglas tubs located in a temperature- and humidity-controlled vivarium. Behavioral training and testing was conducted during the light phase of the 12 h light/dark cycle. Rats were fed 10–15 g of home chow after each daily training session, which was sufficient to maintain them at ~90% of their free-feeding body weight. Rats had *ad libitum* access to water in their home cage. All procedures were approved by the University of California, Los Angeles Animal Research Committee.

Behavioral testing took place in 24 operant chambers enclosed in sound- and light-attenuating shells (Med Associates). Each chamber was equipped with two retractable levers that could be extended to the left and right of a recessed food magazine. Attached to the food magazine were two pellet dispensers that were used to deliver a 45 mg grain or sucrose food pellet (Bioserv Biotechnologies) and two syringe pumps used to deliver 0.1 ml of either 20% polycose or 20% sucrose solution. An infrared photobeam crossed the magazine opening, allowing for the detection of head entries. Illumination of the operant chamber was provided by a 3 W, 24 V house light. A set of three microcomputers running Med Associates proprietary software (Med-PC) controlled all experimental events and recorded lever presses and magazine entries.

Behavioral procedures: instrumental training and testing
Master versus yoked. Before instrumental training, rats were trained to approach the food magazine, in which they were confined to the operant chamber for 30 min while 20 mg sucrose pellets were delivered into the food magazine on a random-time 60 s schedule. The day after magazine training, rats were given one session in which a single lever was inserted into the chamber. For the master group, every lever press delivered a single 20 mg sucrose pellet. Yoked controls received the same temporal pattern of reward delivery that occurred independent of their activity on the lever. Sessions were terminated after 20 outcomes were delivered. After completion of the training session, rats were removed from the chamber and treated for Western blot as described below.

Acquisition versus additional training. After magazine training, rats were given one session on a continuous reinforcement schedule, as described above, after which they underwent additional instrumental training under a variable interval (VI) schedule of reinforcement. Under the VI schedule, a reinforcer is available after a random period of time drawn from a distribution of times centered on 15, 30, or 60 s. Rats received 2 d of VI-15 training, 2 d of VI-30 training, and were killed after either 1 or 5 d of VI-60 training and treated for Western blot as described below.

Inhibition of ERK activity in pDMS and DLS during training. Before instrumental training, rats were trained to approach the food magazine,

in which they were confined to the operant chamber for 30 min while 0.1 ml of 20% polycose solution was delivered into the food magazine on a random-time 60 s schedule. Rats experienced one magazine training session per day for 3 d. After magazine training, rats were trained in separate sessions to respond under a random ratio (RR) schedule on either the left or right lever for 20% polycose delivery, and the order of lever training (left or right) was randomized across subjects and alternated across training days. For the first 2 d, each response the rat made on the lever resulted in response-contingent reward (CRF), followed by 2 d in which outcomes were delivered according to a RR-5 schedule, in which on average five responses resulted in polycose delivery, followed by 2 d under an RR-10 schedule. Each session was 20 min in duration.

After instrumental training for polycose delivery, rats underwent cannulation surgery. After recovery from surgery, rats were trained to associate actions on each of the two levers with a unique outcome. For half of the subjects in each group, pressing the left lever delivered grain pellets and pressing the right lever delivered sucrose pellets, whereas the remaining subjects received the opposite action–outcome relationship. During the training session, one of the two levers was available, and outcomes were delivered according to an RR-10 schedule. Rats received one session with each lever, and each training session lasted 15 min with a 5 min interval between sessions. The MAP kinase kinase (MEK)/ERK inhibitor U0126 [1,4-diamino-2,3-dicyano-1,4-bis(*o*-aminophenylmercapto)butadiene] (Millipore Corporation) or vehicle infusions occurred 30 min before the first acquisition session. The next day, one of the two outcomes was devalued by specific satiety, and rats' choice performance was tested. Rats were provided with 20 g of one of the two outcomes in their home cage for 1 h. Rats were then placed in the operant chamber, the house light was illuminated, and both levers were inserted. Rats could respond on either of the two levers, which produced no outcomes during the session. The session terminated after 10 min, and rats were returned to their home cage.

Inhibition of ERK activity in pDMS and DLS on test. Identical procedures were used for instrumental acquisition as those described above, except that rats received an additional 2 d (for a total of 3 d) of instrumental training with two outcomes. Eight rats from the previous experiment (pretraining U0126 or vehicle infusions) were used for pretest infusions and infusates, and choice of outcome for devaluation were counterbalanced across experiments for these animals. After instrumental training, rats' choice performance was tested after outcome devaluation using the procedures described above. Infusions of U0126 or vehicle occurred 30 min before the choice test.

Surgery and intracranial infusion
Rats were anesthetized with an intraperitoneal injection of sodium pentobarbital (50 mg/kg body weight). Rats were placed in a stereotaxic instrument, the scalp was retracted, and two small holes were drilled in the skull at the following coordinates relative to bregma (in mm): pDMS, −0.4 anteroposterior (AP), ±2.6 mediolateral (ML), −4.5 dorsoventral (DV); DLS, +0.7 AP, ±3.6 ML, −4.5 DV. Two 26 gauge guide cannulae that extended 4.5 mm below the mounting pedestal were lowered through the holes and affixed to the skull with dental cement. Two skull screws were placed posterior to the cannula implantation site and anchored the dental cement. Rats were allowed 1 week of recovery after surgery.

During infusion, rats were restrained by hand, and a 32 gauge infusion cannula was inserted into the guide cannula. The infusion cannula extended 0.5 mm beyond the tip of the guide cannula and was attached at the other end to a 10 μl Hamilton syringe. The MEK/ERK inhibitor U0126 was diluted in 50% DMSO and 0.8% saline to a concentration of 1 mg/ml. Infusate (0.5 μl) was delivered over the course of 1 min, and the infusion cannula was left in place for an additional minute. Rats remained in their home cage for 30 min after infusion before any behavioral procedure took place.

Western blot
Groups of rats were killed immediately after the cessation of their first CRF training session, their first VI-60 training session, or their fifth VI-60

training session. Brains were rapidly removed and immersed in ice-cold PBS for 30–60 s. Brains were blocked, and ~2 mm^3 of dorsomedial and dorsolateral striatum were hand dissected from each region with the center of the volume being located relative to bregma (in mm): +0.7 AP, ±2.0 ML, 4.4 DV (pDMS) and +0.7 AP, ±3.5 ML, 4.4 DV (DLS). Samples were immediately homogenized in 200 μl of lauryl sulfate (SDS) buffer (100 mM Tris-Cl, pH 6.8, 4% SDS, with the addition of 1:100 protease and phosphatase inhibitor cocktails; Sigma P2850, P5726, and P8340) and stored on ice. Later, samples were centrifuged at 13,000 rpm, and the supernatant was removed for preparation of running samples. Protein was quantified using the Pierce BCA method. The samples were diluted to 10 μg/μl with sample buffer (150 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 100 mM dithiothreitol, and 0.1% bromphenol blue) to produce running samples that were heated 100°C for 10 min.

Protein (100 μg) was loaded in 10% precast acrylamide mini-gels (Bio-Rad) and run at 150 V for 80 min. Protein was transferred to a nitrocellulose membrane using 25 V for 45 min in a semidry transfer cell. Membranes were blocked in a solution of 5% nonfat dry milk and 0.1% Tween 20 in TBS for 1–2 h at room temperature and then incubated in anti-pERK (9101S; Cell Signaling Technology) at 1:1000 for 12 h at 4°C. After washing, membranes were incubated in a goat anti-rabbit HRP-conjugated secondary antibody (170-6515; Bio-Rad) for 1 h at 1:2000 at room temperature. Membranes were visualized using chemifluorescence (Supersignal West Pico; Pierce). Blots were exposed to film and scanned. After being imaged with pERK, membranes were stripped (Restore; Pierce) and reprobed with total ERK (tERK) (9102; Cell Signaling Technology).

Quantification of bands was performed using ImageQuant (GE Healthcare). All blots were treated together and exposed to film in a single case to ensure equal exposures.

Histology
Rats were deeply anesthetized with sodium pentobarbital and perfused transcardially with 0.1 M PBS, followed by a solution containing 10% Formalin diluted in PBS. The brain was removed and placed in a 20% sucrose solution overnight, after which 40 μm coronal sections were mounted on slides and stained with thionin using standard histological techniques. Cannula placement was assessed based on the position of tracts in the brain.

Statistical analysis
All statistical analyses were performed with SPSS (SPSS Inc.). Because of minimal expression of ERK1 in our samples, we confined our analysis of Western blot data to ERK2. Immunoblots were quantified in terms of optical density units (odu). To compute ERK activation, odu values derived from pERK2 immunoblots were expressed as a percentage of tERK2 immunoblot odu values. Two-factor ANOVAs with between- and within-subject factors were used to determine significance with ERK activation as the dependent variable. To analyze instrumental response data in the intracranial infusion experiments, response rates on the two levers were compared across infusate conditions using a two-factor ANOVA with both within-subjects and between-subjects factors. *Post hoc* comparisons were made using paired t tests with Bonferroni's correction for multiple comparisons.

Results
Instrumental learning and performance increases ERK2 activation in dorsal striatum
The hypothesis that ERK2 activation in pDMS is critical for action–outcome encoding in instrumental conditioning implies that we should be able to observe evidence of this activation during the early stages of instrumental training. Unfortunately, no information currently exists to support this assumption. As such, we began this series by first conducting two experiments to assess ERK2 activation in dorsal striatum during the early stages of action acquisition and in subsequent performance in instrumental conditioning. Groups of rats experienced either response-

contingent outcomes or were given outcomes independent of their actions and delivered according to their yoked counterparts. After this experience, ERK activation was probed in these animals. Because of the low levels of pERK1 expression in our immunoblots, we restricted our analysis to ERK2. We found that ERK2 activation (pERK2 as a percentage of tERK2 expression) in both the pDMS and DLS was significantly greater in the response-contingent group than in the yoked controls (Fig. 1A). A two-factor ANOVA conducted on these data revealed a main effect of training group ($F_{(1,14)} = 6.567$, $p < 0.05$) and a main effect of brain region ($F_{(1,14)} = 6.751$, $p < 0.05$) but no interaction between these factors ($p > 0.1$) (Fig. 1C). Independent samples t tests revealed significantly greater ERK2 activation in the response-contingent group compared with the yoked control in the pDMS ($t_{(14)} = 2.54$, $p < 0.05$) and DLS ($t_{(13)} = 2.31$, $p < 0.05$). tERK2 odu values were analyzed separately (supplemental Table 1, available at www.jneurosci.org as supplemental material). No effect of group or region was observed in the expression of tERK1 or tERK2 ($p > 0.05$), indicating that the effect of instrumental training on ERK2 activation was likely the result of increased phosphorylation of ERK2, not differences in the overall amount of ERK2 protein available for phosphorylation.

We also examined pERK expression in response-contingent and yoked control animals using pERK immunohistochemistry (for details on the staining procedure, see supplemental methods, available at www.jneurosci.org as supplemental material). Although the small sample size limits the conclusions we can draw from these data, the pattern of results is similar to that observed using Western blot analysis, i.e., greater expression in DMS and DLS in animals experiencing response-contingent outcomes compared with yoked controls (supplemental Fig. S1, available at www.jneurosci.org as supplemental material).

In a second experiment, we assessed the effects of extended instrumental training on ERK2 activation in the dorsal striatum by comparing ERK2 activation in rats given initial acquisition on the lever, as in the first study, plus 1 d of training on a VI-60 schedule of reinforcement (short training group), with rats given four additional training sessions on the VI-60 schedule (extended training group). Confirming our previous result, we found that initial instrumental acquisition activated ERK2 in both the pDMS and DLS, as shown by the level of ERK2 activation observed in the short-training group. Furthermore, extended instrumental training resulted in a similar level of ERK2 activation in the pDMS as was observed during initial training. In contrast, ERK2 activation in the DLS was reduced after extended instrumental training compared with the level observed during initial training. In accord with this description, a two-factor ANOVA conducted on these data revealed a significant effect of striatal subregion ($F_{(1,13)} = 13.97$, $p < 0.05$) and, more importantly, a significant subregion × training group interaction ($F_{(1,13)} = 12.78$, $p < 0.05$) (Fig. 1E). No main effect of training was observed in these conditions ($p < 0.1$). An independent samples t test revealed significantly greater ERK2 activation in the DLS of the short-training group compared with the extended training group ($t_{(13)} = 2.90$, $p < 0.05$). These data show that an increase in ERK2 activation occurs in the pDMS during initial instrumental acquisition and increases with additional training, consistent with the suggestion that ERK2-mediated plasticity underlies the encoding of the action–outcome associations in that region. We next examined this hypothesis by manipulating ERK activation in pDMS during instrumental acquisition.
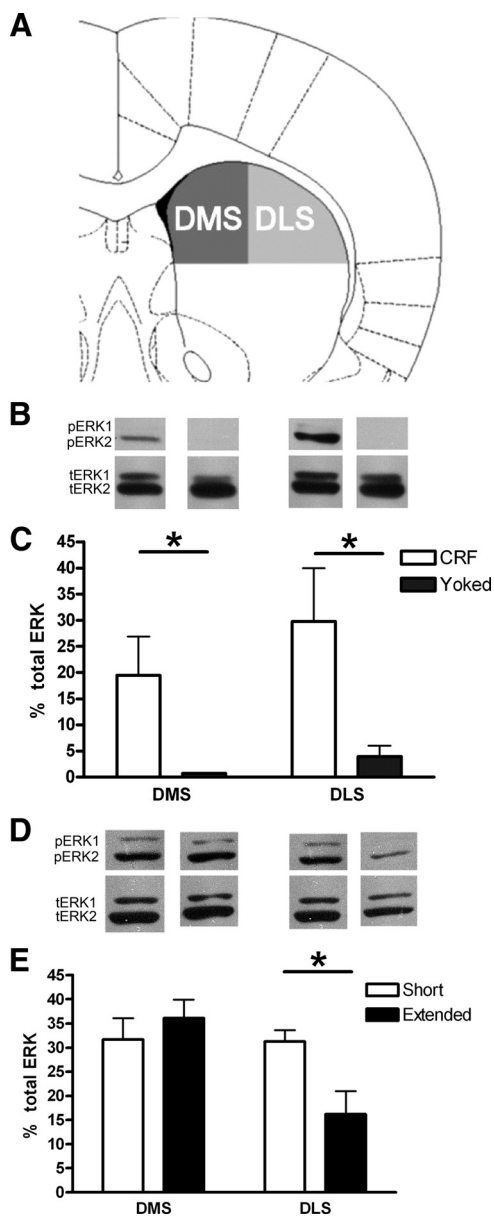
Figure 1. Instrumental learning and performance activates ERK2 in striatal regions. *A*, Schematic depicts locations in which tissue samples from the DMS and DLS were harvested for Western blot analysis. *B*, Example Western blots showing phosphorylated ERK (pERK1 and pERK2) and reference total ERK (tERK1 and tERK2) from the DMS (left two bands) and DLS (right two bands). The bottom (p42) band was used for analysis. *C*, The effect of initial training using response-contingent reward (CRF) and non-contingent reward (YOKED) on ERK2 activation in the DMS and DLS is depicted. An increase in ERK2 activation was observed in both striatal subregions in instrumentally trained animals but not yoked controls. *D*, Example Western blots showing pERK and tERK from the DMS (left two bands) and DLS (right two bands). The bottom (p42) band was used for analysis. *E*, The effect of extended instrumental training on ERK2 activation in the dorsal striatum is depicted. Similar levels of ERK2 activation were observed in the DMS after short instrumental training (1 session) and after extended instrumental training (4 sessions). In contrast, a reduction in the level of ERK2 activation was observed in the DLS after extended instrumental training compared with the level observed after short training. $*p < 0.05$, independent samples t test. Scale bar, ± 1 SEM. Schematic adapted from Paxinos and Watson (2005).

## Disruption of ERK activation in the pDMS prevents the encoding of action–outcome associations

To assess more directly the functional effects of ERK activation in the pDMS during instrumental learning, rats were first trained to respond on each of the two levers (R1 and R2) for a common
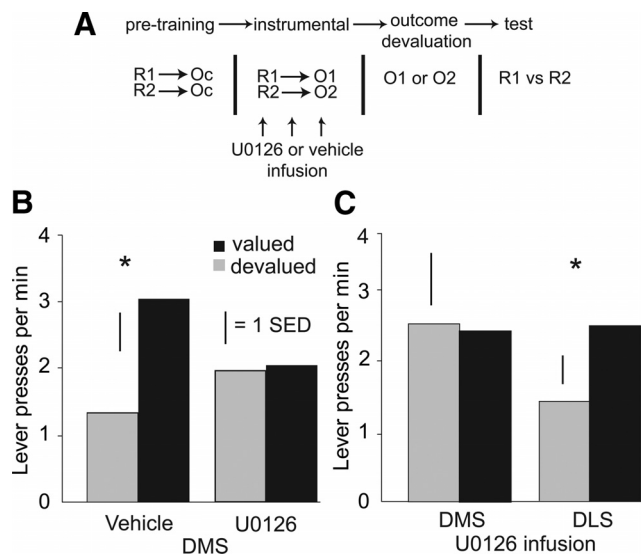
Figure 2. pDMS ERK inactivation prevents the encoding of action–outcome associations during instrumental learning. *A*, The schematic illustrates the experimental procedure. Rats with chronic indwelling cannulae targeting the pDMS underwent a single instrumental conditioning session with actions on each lever associated with a unique outcome. Before this training session, rats received infusions of the MEK/ERK inhibitor U0126 ($n = 9$) or vehicle ($n = 8$). One of the two instrumental outcomes was then devalued with a specific satiety procedure before a 10 min extinction test. *B*, The mean response rate per minute is plotted for the lever that delivered the now devalued outcome or the valued outcome during training for rats that received vehicle or U0126 infusions into the pDMS. Whereas vehicle-infused rats demonstrated sensitivity to outcome value, as shown by a significant reduction in response rates on the lever associated with the devalued outcome, no outcome sensitivity was observed in U0126-infused rats. *C*, Rats received U0126 infusions into the pDMS ($n = 6$) or DLS ($n = 6$) before undergoing a single instrumental session with actions on each lever associated with a unique outcome. Performance on the extinction test after selective devaluation of one of the outcomes revealed sensitivity to outcome value in DLS-infused rats, as shown by a significant reduction in mean response rates on the lever associated with the now devalued outcome compared with the valued outcome. No outcome sensitivity was observed in pDMS-infused rats. Scale bar, $+1$ standard error of the difference (SED), a measure of within-subject variability. $*p < 0.05$, paired t test.

outcome (Oc; a 20% polycose solution). After several training sessions with increasing response requirements for polycose delivery (see Materials and Methods), rats underwent separate instrumental training sessions in which a different type of outcome (O1 and O2, i.e., sucrose or grain pellet) was assigned to each lever (compare Fig. 2*A*). Before each of these training sessions, rats received intra-pDMS infusions of either the MEK/ERK inhibitor U0126 ($n = 9$) or vehicle ($n = 8$). This design allowed for the assessment of ERK inactivation during action–outcome encoding independent from its effects on other forms of learning that might occur during initial instrumental acquisition. Rats that received U0126 or vehicle infusions responded at similar rates on both levers during the training session that followed infusion (mean responses per minute: vehicle, 4.85 ± 1.27; U0126, 5.10 ± 0.87; $p > 0.1$).

The following day, we assessed the effects of the U0126 infusion on encoding the specific action–outcome associations during training using an outcome devaluation test procedure. To achieve this, one of the two instrumental outcomes was devalued using a specific satiety treatment, i.e., the rats were fed one of these two outcomes for 1 h, after which instrumental performance was assessed in an extinction test with both levers available. Rats that received vehicle infusions demonstrated that they had encoded the specific action–outcome associations during the training session and reduced responding on the lever that, during
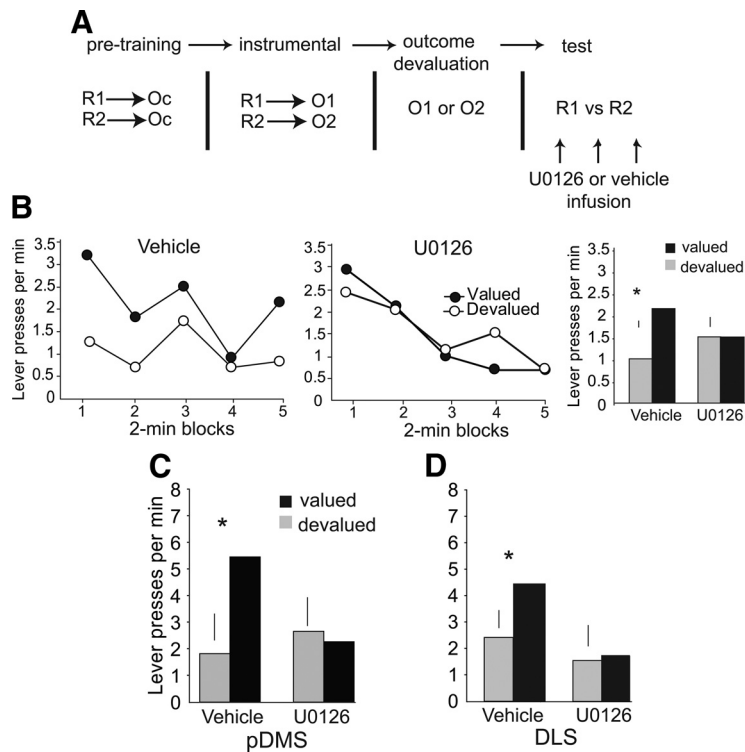
Figure 3. ERK inactivation disrupts instrumental performance guided by action–outcome associations. ***A***, The schematic illustrates the experimental procedure. Rats received U0126 or vehicle infusions into the pDMS or DLS before their performance on the extinction test, in which one outcome was devalued through selective satiety. ***B***, The mean response rates during the 10 min extinction session are plotted for both levers in 2 min bins for pDMS U0126-infused rats ($n = 8$) and vehicle-infused rats ($n = 7$). Whereas vehicle-infused rats demonstrated sensitivity to outcome value, as shown by a significant reduction in response rates on the lever associated with the devalued outcome, no outcome sensitivity was observed in pDMS U0126-infused rats. The bar graph represents the average response rate for the 10 min test session for each group. ***C***, ***D***, pDMS and DLS ERK inactivation disrupts instrumental performance guided by action–outcome associations. Rats received U0126 or vehicle infusions into the pDMS ($n = 6$) or DLS ($n = 5$) before their performance on an extinction test with one of the instrumental outcomes devalued. ***C***, After vehicle infusion, pDMS-infused rats show a reduction in response rates on the lever associated with the devalued outcome; however, no sensitivity to outcome value was observed in rats after pDMS U0126 infusion. ***D***, After vehicle infusion, DLS-infused rats show a reduction in response rates on the lever associated with the devalued outcome; however, no sensitivity to outcome value was observed in rats after pDMS U0126 infusion. dev, Devalued; val, valued. Scale bar, +1 SED. $*p < 0.05$, paired t test.

training, had delivered the now devalued outcome. In contrast, instrumental performance in the rats that received U0126 infusions was insensitive to the devaluation treatment, and they showed no difference in the rate of responding between the two levers (Fig. 2*B*). A two-factor ANOVA with a within-subjects factor of devaluation and a between-subjects factor of drug treatment revealed a significant interaction between these factors ($F_{(1,15)} = 5.30$, $p < 0.05$). *Post hoc* comparisons revealed that vehicle-infused rats performed significantly fewer presses on the lever that, during training, had delivered the now devalued outcome (paired t test, $t_{(7)} = 3.64$, $p < 0.01$), whereas U0126-infused rats showed no difference in response rates on the two levers ($p > 0.1$).

Next we assessed the specificity of the effects of U0126 infusion into the pDMS on action–outcome encoding, comparing the effects of an infusion made into the pDMS with those of an infusion made into the DLS. Twelve rats (six per group) with guide cannulae targeting the pDMS or the DLS received U0126 infusions before being presented with novel action–outcome pairings exactly as described previously (compare Fig. 2*A*). The response rate during training did not differ between rats that received pDMS and DLS U0126 infusions (responses per minute: pDMS, 4.53 ± 0.69; DLS, 5.14 ± 0.91; $p > 0.1$). In agreement

with our previous results, rats that received pDMS U0126 infusions during training showed no sensitivity to outcome devaluation after specific satiety treatment, responding equally on both levers (Fig. 2*C*). In contrast, rats that received U0126 infusions into the DLS during training showed sensitivity to outcome value, performing fewer presses on the lever that, during training, delivered the now devalued outcome (ANOVA: significant interaction between outcome value and infusion location, $F_{(1,10)} = 5.07$, $p < 0.05$; paired t test: DLS infused rats, devalued vs valued lever $t_{(5)} = 3.92$, $p < 0.05$). Thus, action–outcome association formation depends on ERK activation in the pDMS but not adjacent DLS.

Disruption of ERK activation in both pDMS and DLS disrupts instrumental action selection

In addition to its role in acquisition, the pDMS has been shown to be necessary for expression of action–outcome associations (Yin et al., 2005b). Given these reports, we assessed the role of ERK activation in the pDMS during the performance of actions guided by previously acquired action–outcome associations. After rats were trained to respond on each lever for polycose, they received 3 d of instrumental training with responses on the left and right lever associated with different types of outcomes (sucrose or grain pellet). Instrumental responding was then tested after devaluation of one of the two instrumental outcomes (Fig. 3*A*). In a first experiment assessing the involvement of ERK activation during action selection, two groups of rats received either intra-pDMS U0126 or vehicle infusions before the outcome devaluation test. Rats that received vehicle infusions before the test showed a clear outcome devaluation effect and made fewer responses on the lever that, during training, had delivered the now devalued outcome. In contrast, no sensitivity to outcome devaluation was observed in rats that received U0126 infusions before the test (Fig. 3*B*). A two-factor ANOVA revealed a significant interaction between factors of devaluation and drug treatment on lever response rate ($F_{(1,12)} = 5.91$, $p < 0.05$). Instrumental response rates were significantly higher for the lever associated with the valued outcome compared with the devalued outcome in the vehicle-infused rats ($t_{(6)} = 3.51$, $p < 0.05$.) and did not differ in U0126-infused rats ($p > 0.1$).
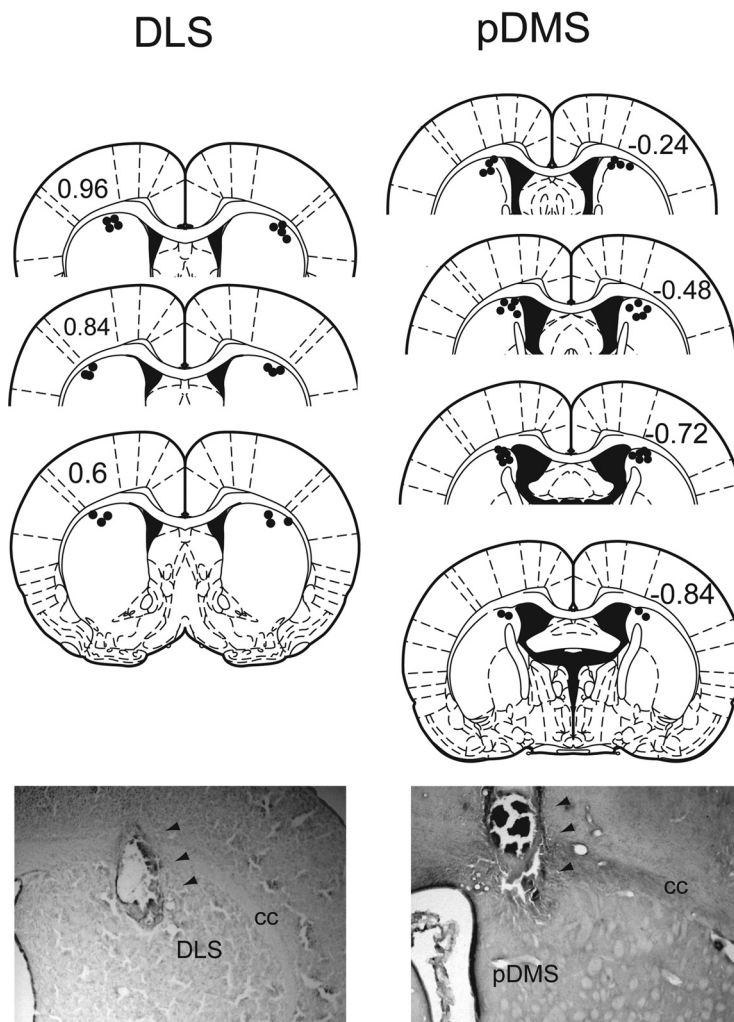
One explanation of the effect of U0126 on performance of previously acquired action–outcome associations is the role of ERK activation not only in synaptic plasticity but in cellular excitability (Yuan et al., 2002; Schrader et al., 2006). On this account, however, ERK activation in structures not directly involved in action–outcome encoding but that are, nevertheless, involved in the performance of instrumental actions should also be affected by these infusions. For example, based on a number of considerations, we have recently argued that the DLS plays a central role in

instrumental performance by controlling a stimulus-based, action-selection process that behavioral research suggests regulates the initiation of actions based on the action–outcome association (Dickinson and Balleine, 1993; Balleine and Dickinson, 1998b; Balleine and Ostlund, 2007; Balleine et al., 2009).

To assess this suggestion, we conducted a second study, based on that described above (Fig. 3A), in which we sought to compare the effects of ERK inactivation in the pDMS with the DLS during performance of goal-directed actions. Eleven rats with cannulae targeting the pDMS ($n = 6$) or DLS ($n = 5$) received 3 d of instrumental training with actions on the left and right lever associated with unique outcomes. Instrumental responding was then tested after devaluation of one of the two instrumental outcomes. Before the test, rats received U0126 or vehicle infusions. They then received an additional instrumental training session followed by a second test preceded by the opposite drug treatment, i.e., if previously given U0126, they were given a vehicle infusion and vice versa. The order of infusate (U0126 → vehicle; vehicle → U0126) was counterbalanced across subjects.

We found that U0126 infusion before test, whether made into the pDMS or the DLS, was effective in abolishing sensitivity to outcome devaluation in this choice situation (Fig. 3C,D). A three-factor ANOVA with drug condition and outcome devaluation as within-subjects factors and infusion location as a between-subjects factor was performed on instrumental response rates in the extinction

Figure 4. Schematic representation of cannulae placement for rats that received infusions into the DLS and pDMS. The numbers on each illustration indicate millimeters from bregma. Adapted from Paxinos and Watson (2005). Photomicrographs from coronal sections depict cannula tracks (marked by arrows) within the DLS (left) and pDMS (right). cc, Corpus callosum.

tests. This analysis revealed a significant interaction between outcome value and drug condition, indicating that, independent of infusion location, rats made fewer responses for the devalued outcome after vehicle infusion but not after U0126 infusion (ANOVA: outcome value $\times$ infusate, $F_{(1,10)} = 8.64$, $p < 0.05$). However, there was no effect of infusion location as a main effect, nor any significant interaction involving this factor (p values >0.05), which suggests that both pDMS and DLS U0126 infusions impair performance during the devaluation test. Instrumental response rates were significantly higher for the lever associated with the valued outcome compared with the devalued outcome in rats that received vehicle infusions into the pDMS ($t_{(5)} = 3.16, p < 0.05$) and DLS ($t_{(4)} = 2.73, p < 0.05$) and did not differ in U0126-infused rats ($p > 0.05$).

Cannulae placement was examined after behavioral testing and is depicted in Figure 4. All animals exhibited visible cannulae tracts extending into striatum. Overall, the effect of U0126 infusions into the dorsal striatum indicate that, in addition to its role in the acquisition of action–outcome associations, pDMS ERK activation is also required for performance guided by previously acquired action–outcome associations. Furthermore, consistent with previous suggestions based on previous assessments of the role of the DLS in stimulus-based action selection, it appears that

ERK activation in the DLS is also required to maintain choice performance after outcome devaluation. Although the DLS is clearly not required for the acquisition of goal-directed actions, its involvement in choice after outcome devaluation suggests that networks involving the pDMS and the DLS work synergistically to control the initiation of actions during instrumental performance.

Discussion

Our results demonstrate a functional role for ERK activation in the dorsal striatum during both the acquisition and performance of goal-directed instrumental actions. The dorsal striatum is functionally segregated into a medial area that is necessary for the encoding and performance of goal-directed actions and a lateral area that is involved in the stimulus control of actions, both in terms of action-selection and performance (Yin et al., 2004; Balleine and Ostlund, 2007; Balleine et al., 2009). We observed an increase in ERK2 activation in both the pDMS and DLS during instrumental acquisition, suggesting that these regions may routinely work in parallel to support the performance of goal-directed actions. We further found that disrupting ERK activation in the pDMS prevented both the encoding of action–outcome associations and the performance of instrumental actions guided by these associations. In contrast, instrumental

acquisition was unimpaired by DLS ERK inactivation; only with additional training did instrumental performance on test become dependent on ERK activation in DLS. These results highlight the dynamic process through which ERK signaling within different subregions of the striatum regulates instrumental learning and performance.

ERK activation in the pDMS mediates instrumental action–outcome learning

Previous research has demonstrated an important role for striatal ERK activation in synaptic plasticity and the development of drug-related behavior. ERK activation is required for the induction of LTP at corticostriatal synapses (Sgambato et al., 1998; Mazzucchelli et al., 2002). Exposure to drugs of abuse increases ERK activation in striatal neurons, and inhibition of ERK activation prevents some of the enduring behavioral changes caused by drug exposure, such as psychomotor sensitization and drug CPP (Valjent et al., 2000; Gerdjikov et al., 2004; Valjent et al., 2004, 2005). Furthermore, overexpression of ERK2 in striatum enhances corticostriatal LTP, memory retention in active and passive avoidance tasks, and expression of drug CPP (Mazzucchelli et al., 2002; Ferguson et al., 2006), which suggests ERK2 activation in the striatum may play a key role in instrumental learning.

Our results demonstrate that ERK2 activation in the pDMS is required for instrumental action–outcome learning. We found an increase in ERK2 activation in the pDMS and DLS when rewarding outcomes are delivered contingent on an instrumental response and that inhibition of ERK activation before instrumental learning prevented the acquisition of specific action–outcome associations. Because of minimal pERK1 expression in our samples, we were unable to draw any conclusions regarding the involvement of this ERK isoform in instrumental acquisition. Recently, pERK1 has been shown to regulate ERK2 activation (Mazzucchelli et al., 2002; Ferguson et al., 2006). It is possible that regulation of ERK2 activation during learning and/or differences in activation across striatal subregions may exist as a result of instrumental learning, and these possibilities remain to be addressed.

ERK activation in the pDMS and DLS mediates instrumental performance guided by action–outcome associations

In addition to its role in synaptic plasticity, ERK activation modulates intrinsic excitability of individual neurons, by targeting the A-type voltage-gated potassium (K^+) channel Kv4.2 (Adams et al., 2000; Yuan et al., 2002; Schrader et al., 2006). In striatum, modulation of K^+ conductance increases the probability of a neuron generating an action potential in response to excitatory synaptic inputs (Mahon et al., 2004). It follows that ERK inactivation by U0126 may render striatal neurons less responsive to excitatory synaptic inputs, such as those carrying information about the current value of a particular outcome (e.g., from basolateral amygdala) (Balleine et al., 2003; Balleine and Killcross, 2006).

The effects of ERK activation on striatal cell excitability may explain our finding that U0126 infusion into the pDMS and DLS prevented rats from adjusting their instrumental responses to a devalued outcome during the choice test. Similar disruptions of the expression of a learned behavior have been observed after ERK inhibition. For example, systemic injection of the ERK inhibitor SL327 (α-[amino[(4-aminophenyl)thio]methylene]-2-(trifluoromethyl)benzeneacetonitrile) prevents the expression of a conditioned locomotor response to a drug-associated context (Valjent et al., 2006). Likewise, U0126 infusion into the nucleus

accumbens prevents the excitatory effect of pavlovian cues on instrumental behavior and the expression of a drug conditioned place preference (Miller and Marshall, 2005; Shiflett et al., 2008). U0126 infusion into the central nucleus of the amygdala (CeN) prevents drug-seeking behavior, and elevated ERK activation in the CeN in response to drug-cue presentation correlates with the incubation of drug craving (Lu et al., 2005; Li et al., 2008). The effects of U0126 infusion before the choice test in our experiment likely do not reflect a simple motor deficit, because infusions did not impair the rats' ability to make rewarded responses when infused before training. Rather, our data are consistent with the notion that striatal ERK activation, through its modulation of cell excitability, enables the flexible expression of learned behaviors.

In contrast to the pDMS, we observed a reduction in ERK2 activation in the DLS with continued instrumental training, which may be attributable to the fact that performance of well rehearsed actions require less neural activity in DLS compared with recently acquired actions (Tang et al., 2007). However, this does not imply that ERK activation within the DLS becomes less relevant in controlling actions with additional training. Indeed, we found that DLS ERK inhibition disrupted choice performance during the devaluation test. It is unknown whether similar patterns of ERK activation occur in the dorsal striatum under training and test situations, which raises the possibility that the disruptive effect of U0126 before test is unrelated to the changes in ERK activation we observed during training. Future studies may address this possibility by examining dorsal striatal ERK activation after specific satiety devaluation and choice behavior.

Contributions of pDMS and DLS to the control of instrumental performance

Control of instrumental performance can be mediated by distinct corticostriatal circuits (Balleine, 2005; Balleine and Ostlund, 2007; Balleine et al., 2007). Whereas acquisition of goal-directed actions is mediated by association areas of the prefrontal cortex and their projections to the pDMS, the influence of these cortical afferents on instrumental performance is time limited (Ostlund and Balleine, 2005). This is not true of the pDMS; this region is necessary for both acquisition and performance of instrumental actions, and the effect of disrupting ERK activation on both acquisition and test performance in the current study provides additional evidence of this. Goal-directed actions can shift to become more habitual (i.e., outcome insensitive) with continued training, a form of learning based on sensorimotor regions of frontal cortex and their projections to the DLS. Whereas some analyses of instrumental performance have argued that the action controllers subserving goal-directed and habitual actions are independent, previous analyses of goal-directed action have allocated an important role to a response selection mechanism mediated by a sensorimotor learning process likely mediated by the DLS (Dickinson and Balleine, 1993; Dickinson and Balleine, 2002; Daw et al., 2005; Balleine and Ostlund, 2007; Everitt et al., 2008; Balleine et al., 2009). Consistent with this claim, although we found no evidence that ERK inactivation in the DLS plays a role in the formation of the action–outcome association, ERK inactivation in the DLS before the retrieval test was found to disrupt the expression of these associations in performance.

It appears clear, therefore, that the performance of goal-directed actions requires the interaction of circuits involving pDMS and the DLS that can be disrupted by U0126 infusion into either structure. This may be derived from the interaction of a stimulus-based (or stimulus–response) action–selection process with action evaluation driven by the action–outcome associa-

tion. From this perspective, the DLS mediates the selection of an action that then results in the evaluation of that action based on its consequences involving the pDMS and related reward circuitry, before selection will result in action initiation. Given this view, the loss of choice performance after the infusion of U0126 into either DLS or pDMS may be thought to reflect the disruption of this selection–evaluation–initiation sequence (Balleine and Ostlund, 2007).

Conclusions

Generally, our results also contribute to an emerging understanding of the molecular mechanisms that regulate instrumental learning and performance. There is a growing consensus that the control of actions depends on a distributed circuit involving the interaction of several cortical–limbic–basal ganglia loops, within which distinct molecular mechanisms appear to modulate local function to uniquely influence this form of learning and performance. Some of these mechanisms are beginning to be identified, for example the role of GPR6 in the vigor of instrumental performance and of calcium/calmodulin-dependent kinase II in the influence of reward-related cues on action selection (Lobo et al., 2007; Wiltgen et al., 2007). To this, we can add the current evidence that ERK, and presumably the downstream transcription factors that depend on its activation, in the corticostriatal network is critical for both instrumental learning and the excitatory influence that integration of striatal processes mediating action selection and evaluation exert on instrumental performance.

References

Adams JP, Anderson AE, Varga AW, Dineley KT, Cook RG, Pfaffinger PJ, Sweatt JD (2000) The A-type potassium channel Kv4.2 is a substrate for the mitogen-activated protein kinase ERK. J Neurochem 75:2277–2287.

Andrzejewski ME, Sadeghian K, Kelley AE (2004) Central amygdalar and dorsal striatal NMDA receptor involvement in instrumental learning and spontaneous behavior. Behav Neurosci 118:715–729.

Balleine BW (2005) Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol Behav 86:717–730.

Balleine BW, Dickinson A (1998a) The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. Anim Learn Behav 26:46–59.

Balleine BW, Dickinson A (1998b) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419.

Balleine BW, Killcross S (2006) Parallel incentive processing: an integrated view of amygdala function. Trends Neurosci 29:272–279.

Balleine BW, Ostlund SB (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. Ann N Y Acad Sci 1104:147–171.

Balleine BW, Killcross AS, Dickinson A (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. J Neurosci 23:666–675.

Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. J Neurosci 27:8161–8165.

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

Dang MT, Yokoi F, Yin HH, Lovinger DM, Wang Y, Li Y (2006) Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. Proc Natl Acad Sci U S A 103:15254–15259.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.

Dayan P, Balleine BW (2002) Reward, motivation, and reinforcement learning. Neuron 36:285–298.

Di Filippo M, Picconi B, Tantucci M, Ghiglieri V, Bagetta V, Sgobio C, Tozzi A, Parnetti L, Calabresi P (2009) Short-term and long-term plasticity at corticostriatal synapses: implications for learning and memory. Behav Brain Res 199:108–118.

Dickinson A (1994) Instrumental conditioning. In: Animal cognition and learning (Mackintosh NJ, ed), pp 40–79. London: Academic.

Dickinson A, Balleine B (1994) Motivational control of goal-directed action. Anim Learn Behav 22:1–18.

Dickinson A, Balleine B (2002) The role of learning in the operation of motivational systems. In: Steven's handbook of experimental psychology (Pashler H, ed), pp 497–533. New York: Wiley.

Dickinson A, Balleine BW (1993) Actions and responses: the dual psychology of behaviour. In: Spatial representation (Eilan N, McCarthy R, Brewer MW, eds), pp 277–293. Oxford: Basil Blackwell Ltd.

Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW (2008) Review. Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. Philos Trans R Soc Lond B Biol Sci 363:3125–3135.

Faure A, Haberland U, Condé F, El Massioui N (2005) Lesion to the nigrostriatal dopamine system disrupts stimulus–response habit formation. J Neurosci 25:2771–2780.

Ferguson SM, Fasano S, Yang P, Brambilla R, Robinson TE (2006) Knockout of ERK1 enhances cocaine-evoked immediate early gene expression and behavioral plasticity. Neuropsychopharmacology 31:2660–2668.

Gerdjikov TV, Ross GM, Beninger RJ (2004) Place preference induced by nucleus accumbens amphetamine is impaired by antagonists of ERK or p38 MAP kinases in rats. Behav Neurosci 118:740–750.

Li YQ, Li FQ, Wang XY, Wu P, Zhao M, Xu CM, Shaham Y, Lu L (2008) Central amygdala extracellular signal-regulated kinase signaling pathway is critical to incubation of opiate craving. J Neurosci 28:13248–13257.

Lobo MK, Cui Y, Ostlund SB, Balleine BW, Yang XW (2007) Genetic control of instrumental conditioning by striatopallidal neuron-specific S1P receptor Gpr6. Nat Neurosci 10:1395–1397.

Lu L, Hope BT, Dempsey J, Liu SY, Bossert JM, Shaham Y (2005) Central amygdala ERK signaling pathway is critical to incubation of cocaine craving. Nat Neurosci 8:212–219.

Mahon S, Deniau JM, Charpier S (2004) Corticostriatal plasticity: life after the depression. Trends Neurosci 27:460–467.

Mazzucchelli C, Vantaggiato C, Ciamei A, Fasano S, Pakhotin P, Krezel W, Welzl H, Wolfer DP, Pagès G, Valverde O, Marowsky A, Porrazzo A, Orban PC, Maldonado R, Ehrengruber MU, Cestari V, Lipp HP, Chapman PF, Pouysségur J, Brambilla R (2002) Knockout of ERK1 MAP kinase enhances synaptic plasticity in the striatum and facilitates striatal-mediated learning and memory. Neuron 34:807–820.

Miller CA, Marshall JF (2005) Molecular substrates for retrieval and reconsolidation of cocaine-associated contextual memory. Neuron 47:873–884.

Ostlund SB, Balleine BW (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J Neurosci 25:7763–7770.

Ostlund SB, Balleine BW (2007) The contribution of orbitofrontal cortex to action selection. Ann N Y Acad Sci 1121:174–192.

Paxinos G, Watson C (2005) The rat brain in stereotaxic coordinates, Ed 5. Burlington, MA: Elsevier.

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

Schrader LA, Birnbaum SG, Nadin BM, Ren Y, Bui D, Anderson AE, Sweatt JD (2006) ERK/MAPK regulates the Kv4.2 potassium channel by direct phosphorylation of the pore-forming subunit. Am J Physiol Cell Physiol 290:C852–C861.

Sgambato V, Pagès C, Rogard M, Besson MJ, Caboche J (1998) Extracellular signal-regulated kinase (ERK) controls immediate early gene induction on corticostriatal stimulation. J Neurosci 18:8814–8825.

Shiflett MW, Martini RP, Mauna JC, Foster RL, Peet E, Thiels E (2008) Cue-elicited reward-seeking requires extracellular signal-regulated kinase activation in the nucleus accumbens. J Neurosci 28:1434–1443.

Tang C, Pawlak AP, Prokopenko V, West MO (2007) Changes in activity of the striatum during formation of a motor habit. Eur J Neurosci 25:1212–1227.

Valjent E, Corvol JC, Pages C, Besson MJ, Maldonado R, Caboche J (2000) Involvement of the extracellular signal-regulated kinase cascade for cocaine-rewarding properties. J Neurosci 20:8701–8709.

Valjent E, Pagès C, Hervé D, Girault JA, Caboche J (2004) Addictive and non-addictive drugs induce distinct and specific patterns of ERK activation in mouse brain. Eur J Neurosci 19:1826–1836.

Valjent E, Pascoli V, Svenningsson P, Paul S, Enslen H, Corvol JC, Stipanovich A, Caboche J, Lombroso PJ, Nairn AC, Greengard P, Hervé D, Girault JA

(2005) Regulation of a protein phosphatase cascade allows convergent dopamine and glutamate signals to activate ERK in the striatum. Proc Natl Acad Sci U S A 102:491–496.

Valjent E, Corvol JC, Trzaskos JM, Girault JA, Hervé D (2006) Role of the ERK pathway in psychostimulant-induced locomotor sensitization. BMC Neurosci 7:20.

Wiltgen BJ, Law M, Ostlund S, Mayford M, Balleine BW (2007) The influence of Pavlovian cues on instrumental performance is mediated by CaMKII activity in the striatum. Eur J Neurosci 25:2491–2497.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189.

Yin HH, Knowlton BJ, Balleine BW (2005a) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505–512.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005b) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513–523.

Yin HH, Knowlton BJ, Balleine BW (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav Brain Res 166:189–196.

Yuan LL, Adams JP, Swank M, Sweatt JD, Johnston D (2002) Protein kinase modulation of dendritic K^+ channels in hippocampus involves a mitogen-activated protein kinase pathway. J Neurosci 22:4860–4868.