The influence of Pavlovian cues on instrumental performance is mediated by CaMKII activity in the striatum

Brian J. Wiltgen,1 Matthew Law,1 Sean Ostlund,1 Mark Mayford2 and Bernard W. Balleine1
1Department of Psychology and the Brain Research Institute, University of California, Los Angeles, CA 90095, USA
2Department of Cell Biology, Scripps Research Institute, La Jolla, CA 92037, USA

Keywords: incentive learning, motivation, nucleus accumbens, reinforcement, reward

Abstract
Pavlovian cues associated with reward exert a powerful motivational influence on the performance of goal-directed actions. This motivational process depends critically on the ventral striatum, although little is known about the cellular and molecular mechanisms that mediate it. In the current experiments we examined the role of calcium calmodulin-dependent kinase II (CaMKII) by using transgenic mice that express a constitutively active form of this kinase. We found that controlled expression of active CaMKII in the striatum did not affect learning but did impair the motivation of goal-directed actions by Pavlovian cues associated with reward. Mutant mice learned to lever press for reward, remained sensitive to outcome devaluation and contingency degradation manipulations, and were able to acquire Pavlovian responses to cues paired with reward. However, Pavlovian cues were completely unable to motivate lever pressing in mutant mice. This was true even in mice trained with the CaMKII transgene turned off and then tested with it turned on. We were also able to suppress transgene expression in impaired mutants and fully restore the motivational effects of reward cues in these animals. Therefore, the current experiments demonstrate that normal CaMKII activity in the striatum is essential for the motivational effects of reward cues on goal-directed actions.

Introduction
It is well established that animals learn associations between their actions and specific outcomes through a process called instrumental conditioning (Dickinson, 1994; Balleine & Dickinson, 1998). Performance of these actions depends on the learned associations and two distinct motivational processes, one derived from the reward value of the instrumental outcome and the other from the excitory influence of cues that predict reward (Dickinson & Balleine, 1995; Balleine, 2001, 2005). In the latter case, cues acquire significance as predictors of reward through Pavlovian conditioning (Rescorla, 1988) and are able to exert motivational control over instrumental actions leading to the same outcome (Colwill & Rescorla, 1988). This motivational process is referred to as Pavlovian–instrumental transfer because it reflects the interaction of Pavlovian and instrumental learning systems (Rescorla & Solomon, 1967; Lolordo, 1971).

Previous studies have shown that the ventral striatum is important for Pavlovian–instrumental transfer (Corbit et al., 2001; Everitt & Robbins, 2005) although, to date, little is known about the local cellular and molecular mechanisms that mediate this excitatory effect. To address this issue we utilized inducible, transgenic mice that express a constitutively active form of calcium calmodulin-dependent kinase II (CaMKII) (Asp-286) specifically in the striatum that express a constitutively active form of this kinase. We found that controlled expression of active CaMKII in the striatum did not affect learning but did impair the motivation of goal-directed actions by Pavlovian cues associated with reward.

Materials and methods

Subjects
The line of transgenic mice used in this experiment (B21) has been described previously (Mayford et al., 1996). Following repeated backcrossing into the C57BL/6J strain transgene expression was restricted to the striatum. For all experiments, mice were group housed and maintained on a 12 h light : dark cycle. For the behavioral experiments, adult mice were put on a food-deprivation schedule to maintain them at ~85% of their free-feeding weight. The UCLA Animal Research Committee approved this study.
In situ hybridization

Methods for brain tissue preparation and in situ hybridization have been reported previously (Mayford et al., 1995). The mice were briefly anesthetized with halothane, decapitated and the brains quickly removed. Coronal cryostat sections (20 µm) were thaw-mounted onto Fisherbrand Superfrost precleaned microscope slides, fixed in 4% paraformaldehyde and frozen at −70 °C until use. Slices were hybridized to a 35S-labeled antisense DNA probe corresponding to a 30-bp sequence in the CaMKII-Asp-286 transgene (CTTCAGG CAGTCGACGTCCCTCTGCTGCTG). This probe hybridizes specifically to the CaMKII-Asp-286 transgene and does not recognize wild-type (WT) CaMKII. Slides were hybridized overnight at 37 °C and washed twice for 10 min each at room temperature in 2 × standard sodium citrate, then twice for 60 min each at 65 °C. Following exposure to photographic film (Kodak Biomax MR), slides were dipped in Kodak NTB2 emulsion (diluted 1 : 1 in distilled water) at 43 °C, dried flat on ice-cold glass plates and exposed for 5–6 weeks at 4 °C. Slides were developed in Kodak Dektol Developer according to the manufacturer’s instructions, and counterstained in 0.1% thionin.

Behavioral procedures

The conditioning apparatus and procedures have been described previously (Carvalho et al., 2001; Corbit et al., 2001; Ostlund & Balleine, 2005). The mice were trained in eight operant chambers (Medical Associates; East Fairfield, VT, USA) housed in sound- and light-attenuating shells. Each chamber contained two retractable levers, a pellet dispenser that delivered 10 mg Noyes food pellets (A/1 formula) and a dipper that delivered 0.01 mL of a 10% sucrose solution.

Magazine training

Mice received 2 days of magazine training (WT, n = 27; gene on, n = 21). Each day consisted of one 30-min session where pellets and sucrose were delivered on independent random interval (60 s) schedules. The levers were withdrawn during these sessions.

Lever press training

After magazine training, mice were trained to lever press on random ratio (RR) reinforcement schedules. Each day consisted of two separate sessions, one with each lever. For half of the mice the left lever earned pellets and the right lever earned sucrose. This contingency was reversed for the remaining animals. Each session lasted until the animal earned 25 reinforcers or 30 min had expired. The order of the sessions was alternated each day. The time between sessions was at least 1 h. Animals first received six sessions of continuous reinforcement where each press produced a reinforcer. They were then shifted to a RR5 schedule for 4 days, where the probability of reinforcement for each press was 0.2. This was followed by 4 days of a RR10 schedule (probability of reinforcement, P = 0.1) and 4 days of a RR20 schedule (probability of reinforcement, P = 0.05).

Devaluation test

The day after the last RR20 schedule, mice received ad libitum access to one of the reinforcers for 1 h. Half of the animals received pellets and the remaining mice received sucrose. Immediately after access to the reinforcer mice were placed in the operant chamber for a 10-min choice extinction test. The next 2 days the mice were retrained on a RR20 reinforcement schedule and were then given a second devaluation test. This test was identical to the first devaluation test except that animals were given access to the opposite reinforcer. The data were plotted and analysed during the test as a percentage of responding during the last day of RR20 training.

Contingency degradation

Following the second devaluation test mice were retrained on a RR20 schedule for 4 days. Afterwards, they received 4 days of contingency degradation training. During these 20-min sessions, mice continued to receive contingent reinforcement (P = 0.05) for lever pressing. In addition, free outcomes were delivered non-contingently during every second the animals did not press (P = 0.05). During one session the non-contingent outcome was the same as that earned by lever pressing (degraded contingency). During the other session the non-contingent outcome was different than that earned by lever pressing (non-degraded). Therefore, for the degraded response–outcome contingency the reinforcer was delivered at the same probability whether or not the animal pressed. For the non-degraded contingency the reinforcer could only be earned by pressing. Half of the animals had the lever press–pellet contingency degraded and half of the animals had the lever press–sucrose contingency degraded. The day after the last degradation training session the mice received a 10-min choice extinction test. The data were plotted and analysed during degradation training as a percentage of responding during the last day of RR20 training.

Pavlovian conditioning

Starting 2 weeks prior to Pavlovian conditioning, some mice had doxycycline (200 µg/mL; WT, n = 20; mutant, n = 24) added to their water while others received no treatment (WT, n = 12; mutant, n = 16). After this period, mice received 8 days of Pavlovian conditioning. Each day consisted of a single hour-long session in which two conditional stimuli [tone (85 dB, 2000 Hz) or lights off] were paired with reinforcement. The levers were retracted during these sessions. Each stimulus was presented four times during each session. For half of the mice the tone was paired with pellets and dark was paired with sucrose. This relationship was reversed in the remaining animals. Each stimulus was presented for 2 min, during which reinforcement was delivered on a random interval schedule (30 s). The time between stimulus presentations was ~5 min. The number of magazine entries (conditional response) was recorded during the stimuli and during a 2-min prestimulus interval. The data were analysed as an elevation ratio (entries during conditioned stimulus (CS)/entries during CS + prestimulus entries).

Transfer test

After the last day of Pavlovian conditioning mice received a 45-min choice extinction test during which the CS were presented. No reinforcement was delivered during this session. During the first 8 min the levers were available, but no stimuli were presented. Each stimulus was delivered four times and presentations were separated by ~5 min. Lever presses were recorded during the stimuli and during a subsequent 2-min interval. Three groups of animals underwent this test: mice that received doxycycline during training and testing (WT, n = 12; mutant, n = 16), mice that did not receive doxycycline (WT, n = 12; mutant, n = 16) and mice that received doxycycline during training, but nothing during the test (WT, n = 8; mutant,
After the initial transfer test, some of the mutant animals had their treatment conditions switched and were then retrained and retested (gene off/on, \( n = 16 \); gene on/off, \( n = 10 \)).

### Reinstatement test

A reinstatement test was administered to WT (\( n = 24 \)), mutants on doxycycline (\( n = 8 \)) and untreated mutants (\( n = 24 \)) after four additional days of lever press training. Mice were extinguished on both levers for 15 min, after which time a single reinforcer was delivered. Lever presses were recorded during the last 3 min of extinction and the 3 min period after the reinforcer was delivered.

### Statistical analyses

All between-subject comparisons were made using an analysis of variance (ANOVA). Within-group comparisons were made using a
repeated-measures ANOVA. Post hoc tests were done with Fisher’s PLSD. For all tests, \( P < 0.05 \) was used to indicate statistical significance.

Results

**Instrumental learning is normal in CaMKII mutants**

We evaluated the role of CaMKII in the striatum by generating a line of mice that selectively express a mutant form of this kinase (Asp-286) in this region (Fig. 1A–D). Expression was highly selective and not observed in adjacent structures like the amygdala (Fig. 1E and F). To determine the effects of active CaMKII on goal-directed instrumental actions, mutants with the gene on were trained to lever press for reward. Each animal was trained to press two levers, one rewarded with food pellets and another by the delivery of a sucrose solution. Figure 2 shows the acquisition of lever pressing by WT and mutants across training sessions (data averaged from both levers). Both control mice (main effect of session \( F_{3,78} = 448.64, P < 0.05 \)) and mutants with the gene on (main effect of session \( F_{3,60} = 309.25, P < 0.05 \)) showed increases in lever pressing as the number of responses required to receive a reward was increased.

A hallmark of instrumental conditioning is that animals encode the consequences of their actions, i.e. they learn associations between their responses and the specific outcomes that they produce (Dickinson, 1994). This can be demonstrated using a devaluation procedure in which the value of a particular reward is reduced prior to a choice extinction test (Colwill & Rescorla, 1985). We used this test to determine if active CaMKII in the striatum impaired the acquisition of response–outcome associations. In this experiment both WT (main effect of lever \( F_{1,26} = 14.83, P < 0.05 \)) and mutants (main effect of lever \( F_{1,20} = 8.93, P < 0.05 \)) showed a selective reduction in responding on the lever that in training delivered the devalued outcome (Fig. 3A). The size of this reduction did not differ between groups (no lever–group interaction \( F < 1 \)). This result demonstrates that the mutants were able to encode the response–outcome associations to which they were exposed during training. It also suggests that these animals were sensitive to changes in the incentive value of rewarding outcomes (Balleine, 1998).

Instrumental performance is also sensitive to changes in the contingency between a response and the reward it produces. For example, degrading the response–outcome contingency by delivering a particular outcome with a similar probability both when an action is performed and when it is not selectively attenuates the performance of that action. This suggests that animals encode the causal relationship between their actions and the outcomes they produce (Balleine & Dickinson, 1998; Corbit & Balleine, 2000). To determine if this process was impaired by active CaMKII expression in the striatum, animals were given several sessions in which one of the response–outcome contingencies on which they were trained was degraded. The results of these sessions are shown in Fig. 3B. Both WT (main effect of lever \( F_{1,26} = 9.51, P < 0.05 \)) and mutants (main effect of lever \( F_{1,20} = 9.61, P < 0.05 \)) showed a selective reduction in responding on the lever for which the action–outcome contingency had been degraded. The size of this reduction did not differ between groups (no lever–group interaction \( F < 1 \)). To ensure that the reduced responding was not simply the result of outcome-specific satiety, animals received a choice extinction test on the two levers the following day, and both WT (main effect of lever \( F_{1,26} = 19.99, P < 0.05 \)) and mutants (main effect of lever \( F_{1,20} = 13.44, P < 0.05 \)) continued to show less responding on the degraded lever.

**Active CaMKII prevents Pavlovian cues from motivating instrumental actions**

We next conducted Pavlovian conditioning with discrete cues for reward. Prior to these sessions, some animals were treated with doxycycline (200 \( \mu \)g/mL for 2 weeks) to prevent the expression of constitutively active CaMKII. Doxycycline had no effect on WT mice (no effect of treatment \( F < 1 \)), so treated and untreated animals were combined for the analyses. Conditioned magazine entries were acquired by WTs (main effect of day \( F_{7,161} = 3.95, P < 0.05 \)) and mutants (main effect of day \( F_{7,161} = 12.04, P < 0.05 \)) and mutants with the gene on (main effect of day \( F_{7,105} = 12.46, P < 0.05 \)) across training days. At the end of training,
mutants with the transgene off (main effect of condition $P < 0.05$) showed a clear enhancement in responding on the lever that led to the same outcome during training. Mutants with the gene on showed absolutely no change in responding after reward delivery (no effect of condition $P > 0.05$) and mutants with the gene off (main effect of condition $P > 0.05$) and showed no change in instrumental responding relative to baseline (same/different vs. baseline post hoc tests, $P > 0.05$).

To determine if this was a deficit in performance, we also examined Pavlovian–instrumental transfer in a group of mutants trained with the gene off and turned on only for the transfer test. These mutants also showed no influence of the reward-related cues on instrumental performance relative to baseline (same/different vs. baseline Fisher’s PLSD, $P > 0.05$).

To establish whether the effects of the transgene were transient, we switched the two groups of mutant mice with respect to their exposure to doxycycline and retested them. As a consequence, mutants trained with the transgene on were now tested with it turned off (on/off), whereas those trained with it off were now tested with it turned on (off/on). Figure 4B shows the relative change in lever pressing that occurred when the Pavlovian stimulus leading to the same outcome was presented. The data are presented as a difference score for each test (cue-baseline). Turning the transgene off in mutants that were previously impaired restored the excitatory effect of the reward-related stimuli. Conversely, turning the gene on in mutants that previously showed normal sensitivity to reward-related cues now abolished that effect (test-treatment interaction, $F_{1,24} = 7.23$, $P < 0.05$). These results clearly demonstrate that CaMKII activity can be used to control the motivational effects of reward-related cues on goal-directed behavior.

In order to confirm that the transfer deficit in mutants was due to the motivational effect of reward cues, we also used a reinstatement procedure. Reinstatement is a good assay of outcome-mediated initiation of goal-directed behavior in which delivery of the reward itself primes and motivates instrumental actions after a period of extinction (Delamater, 1997; Ostlund & Balleine, 2007). This motivational process has been studied extensively and, like transfer, is known to depend on the ventral striatum (Cornish & Kalivas, 2000; McFarland & Kalivas, 2001). After a period of extinction on both levers, a single reinforcer was delivered to groups of WT and mutant mice (Fig. 5). In WT mice (main effect of condition $F_{2,46} = 3.79$, $P < 0.05$) and mutants with the gene off (main effect of condition $F_{2,14} = 4.43$, $P < 0.05$) this reward strongly motivated instrumental responding on the lever that led to the same outcome during training (same vs. baseline post hoc tests, $P < 0.05$). In contrast, mutant mice with the gene on showed absolutely no change in responding after reward delivery (no effect of condition $F_{2,46} = 1.54$, $P > 0.05$). Hence, mutants expressing constitutively active CaMKII showed a motivational deficit both when the rewarding outcome was delivered and when it was anticipated on the basis of Pavlovian cues.

The motivational effect of reward cues on goal-directed action was then assessed in a test of Pavlovian–instrumental transfer. All mice received a choice extinction test on the two levers during which both of the reward cues were presented intermittently. Doxycycline had no effect on WT mice (no effect of treatment $F_{1,30} = 3.22$, $P > 0.05$) and so, again, treated and untreated animals were combined for the figure and analysis. Figure 4A illustrates that when the cues were presented both WT mice (main effect of condition $F_{2,42} = 5.75$, $P < 0.05$) and mutants with the transgene off (main effect of condition $F_{2,30} = 5.92$, $P < 0.05$) showed a clear enhancement in responding on the lever that, in training, delivered the same outcome as that predicted by the cue (same vs. baseline post hoc tests, $P < 0.05$). This effect is referred to as Pavlovian–instrumental transfer (Rescorla & Solomon, 1967; Loracle, 1971). In contrast, responding on the lever leading to a different outcome did not change relative to baseline (different vs. baseline post hoc tests, $P < 0.05$). Mutants with the transgene on were insensitive to the motivational influence of reward cues and showed no change in instrumental responding relative to baseline (same/different vs. baseline post hoc tests, $P > 0.05$).

In order to confirm that the transfer deficit in mutants was due to the motivational effect of reward cues, we also used a reinstatement procedure. Reinstatement is a good assay of outcome-mediated initiation of goal-directed behavior in which delivery of the reward itself primes and motivates instrumental actions after a period of extinction (Delamater, 1997; Ostlund & Balleine, 2007). This motivational process has been studied extensively and, like transfer, is known to depend on the ventral striatum (Cornish & Kalivas, 2000; McFarland & Kalivas, 2001). After a period of extinction on both levers, a single reinforcer was delivered to groups of WT and mutant mice (Fig. 5). In WT mice (main effect of condition $F_{2,46} = 3.79$, $P < 0.05$) and mutants with the gene off (main effect of condition $F_{2,14} = 4.43$, $P < 0.05$) this reward strongly motivated instrumental responding on the lever that led to the same outcome during training (same vs. baseline post hoc tests, $P < 0.05$). In contrast, mutant mice with the gene on showed absolutely no change in responding after reward delivery (no effect of condition $F_{2,46} = 1.54$, $P > 0.05$). Hence, mutants expressing constitutively active CaMKII showed a motivational deficit both when the rewarding outcome was delivered and when it was anticipated on the basis of Pavlovian cues.

**Discussion**

The current experiments provide clear evidence that CaMKII activity in the striatum is important for the motivational influence of reward-related cues. Expression of a constitutively active form of this kinase prevented Pavlovian cues from augmenting instrumental performance without impairing either Pavlovian or instrumental learning. The deficit observed in mutant mice was not permanent and could be reversed by turning off gene expression with doxycycline. This pattern of results suggests that expression of active CaMKII in the striatum blocks the outcome-selective interaction of Pavlovian and instrumental learning systems, an interaction that previous evidence suggests
involves the nucleus accumbens (Corbit et al., 2001). Consistent with this idea, a recent electrophysiological analysis of our mutant animals found that medium spiny neurons in the shell of the nucleus accumbens exhibit significantly altered firing patterns following active CaMKII expression. Mutant neurons showed no changes in basal transmission or synaptic strength, but did exhibit significantly reduced cellular excitability (Said Kourrich and Mark J. Thomas, personal communication). Similar results have been observed in the rat hippocampus and in Drosophila neurons, suggesting that CaMKII plays a critical role in the regulation of neuronal excitability (Park et al., 2002; Varga et al., 2004). Changes in excitability could produce the deficits observed in our mutants by reducing the responsiveness of accumbens neurons to motivational stimuli during the transfer and reinstatement tests. Similar to our behavioral effects, the reduction in excitability caused by active CaMKII expression can be reversed by turning off gene expression with doxycycline (Said Kourrich and Mark J. Thomas, personal communication).

It is possible that expression of active CaMKII merely alters striatal function generally, much like some form of inactivation, and is not selective to the mechanisms of transfer. Several features of the current results argue against this claim. First, transgene expression is not localized to the ventral striatum but is also expressed in the dorsal striatum, an area where evidence suggests that N-methyl-d-aspartate (NMDA)-mediated neuronal plasticity is important for instrumental learning (Yin et al., 2005). The fact that action–outcome learning appeared normal in mutants with the transgene on suggests therefore that the transgene did not impair plasticity critical to instrumental learning. Furthermore, the relatively specific effect of the transgene on Pavlovian–instrumental transfer also argues against a general effect localized, e.g., to the ventral striatum. In similar fashion to the current study, we have found that selective lesions of the accumbens shell attenuate selective transfer effects, but have no effect whatever on the sensitivity of rats to the selective devaluation of the instrumental outcome by specific satiety (Corbit et al., 2001). This is not true of lesions of accumbens core, however. Although these lesions had no influence on the selective transfer effect abolished by the shell lesions, they had a profound effect on the sensitivity of rats to the selective devaluation of the instrumental outcome (Corbit et al., 2001). The sparing of outcome devaluation in the current series suggests therefore that the accumbens was not generally inactivated. In fact, these data make the further point that, not only are the motivational processes engaged by transfer and devaluation mediated by distinct circuits within the accumbens, they may also be mediated by distinct intracelluar processes. We should note, however, that we did observe transgene expression throughout the dorsal and ventral striatum. Therefore, we cannot rule out the possibility that our behavioral effects arise from changes outside the shell of the accumbens. Future studies will be needed to address this issue and determine the precise locus of our effect within the striatum.

In addition to suggesting a potential cellular mechanism for the interaction of Pavlovian and instrumental learning systems, these data also suggest that the striatum plays a central role in decision making, particularly in the way that predictions of future reward, and the values of associated actions derived thereby, act to determine action selection and initiation (Sutton & Barto, 1998). Previous suggestions along these lines have proposed that the learning of actions and values are mediated by distinct neural structures without indicating how they interact (Daw et al., 2005, 2006). The current data suggest that the interaction of these processes occurs at neurons in the striatum and is regulated by changes in excitability that disrupt the fidelity of firing patterns and the activation of downstream target networks involved in action selection and initiation. Indeed, careful consideration of these data suggests that decisions influenced by exploration/exploitation trade-offs may be more productively viewed in terms of motivational factors rather than the consequence of purely cognitive demands such as uncertainty (Daw et al., 2006), i.e., the tendency to explore may be stronger when motivational demands are relatively weak and the tendency to exploit made more likely as motivational demands increase.

Acknowledgments

We thank Robert Brown and Sheena Josselyn for comments on earlier versions of this manuscript. This research was supported by grants from the National Institutes of Health to B.B. (MH 56446) and M.M. (MH 57368), and a UCLA Research Mentorship Fellowship to B.W.

Abbreviations

CaMKII, calcium calmodulin-dependent kinase II; CS, conditioned stimulus; RR, random ratio; WT, wild-type.

References


