

Brain Circuits of Methamphetamine Place Reinforcement Learning: The Role of the Hippocampus-VTA Loop

Yonas B. Keleta & Joe L. Martinez

University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249 USA

Keywords

Hippocampus, learning and memory, nucleus accumbens, place conditioning, reward, VTA.

Correspondence

Yonas Keleta, Mission Primary Care Medical Group, 3755 Beacon Ave, Fremont, CA 94538, USA. E-mail: yonas.keleta@gmail.com

Supported by the diversity program in neuroscience at NIMH, The Ewing Halsell Distinguished Chair under the American Psychological Program, APA T32 MH18882 and by the MBRS program at University of Texas at San Antonio, NIH GM07717.

Received: 15 August 2011;
Revised: 11 December 2011;
Accepted: 14 December 2011

doi: 10.1002/brb3.35

Abstract

The reinforcing effects of addictive drugs including methamphetamine (METH) involve the midbrain ventral tegmental area (VTA). VTA is primary source of dopamine (DA) to the nucleus accumbens (NAc) and the ventral hippocampus (VHC). These three brain regions are functionally connected through the hippocampal-VTA loop that includes two main neural pathways: the bottom-up pathway and the top-down pathway. In this paper, we take the view that addiction is a learning process. Therefore, we tested the involvement of the hippocampus in reinforcement learning by studying conditioned place preference (CPP) learning by sequentially conditioning each of the three nuclei in either the bottom-up order of conditioning; VTA, then VHC, finally NAc, or the top-down order; VHC, then VTA, finally NAc. Following habituation, the rats underwent experimental modules consisting of two conditioning trials each followed by immediate testing (test 1 and test 2) and two additional tests 24 h (test 3) and/or 1 week following conditioning (test 4). The module was repeated three times for each nucleus. The results showed that METH, but not Ringer's, produced positive CPP following conditioning each brain area in the bottom-up order. In the top-down order, METH, but not Ringer's, produced either an aversive CPP or no learning effect following conditioning each nucleus of interest. In addition, METH place aversion was antagonized by coadministration of the *N*-methyl-D-aspartate (NMDA) receptor antagonist MK801, suggesting that the aversion learning was an NMDA receptor activation-dependent process. We conclude that the hippocampus is a critical structure in the reward circuit and hence suggest that the development of target-specific therapeutics for the control of addiction emphasizes on the hippocampus-VTA top-down connection.

Introduction

Methamphetamine (METH) is one of the most abused psychostimulants in the United States (NIDA report 2006). This nationwide increase in the abuse of METH is believed to be due to its effects on reinforcement learning. The theory of reinforcement learning explains that reward is a stimulus toward which the organism increases the probability of response following the repeated occurrence of the reward and environmental cues paired with it, whereas aversive stimulus decreases the probability of response (Cannon and Palmiter 2003; Rossato et al. 2009). In mammals, including rodents, the rewarding effects of a stimulus can be studied using several behavioral models such as conditioned place preference (CPP) is commonly used to study Pavlovian classical condi-

tioning. Interestingly, CPP is thought to be encoded through the induction of synaptic plasticity including long-term potentiation (LTP) and long-term depression (LTD) (Adamec 2001; Bannerman et al. 2008). Thus, researches in the field of addiction argue that repeated exposure to psychostimulants such as METH results in the long-term alterations of synaptic plasticity in brain areas that are involved in reinforcement learning and reward processing (Kauer and Malenka 2007; Brown et al. 2008).

At cellular level, METH binds to dopamine (DA) transporters, which leads to enhanced DA release through these transporters and thereby increases extracellular levels of DA at cortical and subcortical targets of the ventral tegmental area (VTA). Behavioral electrophysiological investigations argue that the VTA is responsible for encoding of information

relevant to the acquisition phases of positive reinforcement learning (reward) and aversion (Carter and Fibiger 1977). Both the nucleus accumbens (NAc) and the hippocampus receive DAergic innervation from the VTA (Gasbarri et al. 1994, Gasbarri et al. 1997). Functionally, this triad network of these three limbic regions together with the accompanied neurotransmitters and neuromodulators is important not only for enhancing spatial and episodic memories (Broadbent et al. 2004; Ryan et al. 2010), but also for encoding the entry of novel information to the central nervous system (CNS; Jenkins et al. 2004; Lisman and Grace 2005; Lee et al. 2005). Hence, processing of sensory information, whether it is rewarding or aversive hypothetically requires the detection of stimulus novelty or familiarity through the synchronous connectivity of the hippocampus (especially the ventral hippocampus, VHC) and the VTA.

There are two major pathways (routes) in the hippocampus-VTA loop; the top-down route and the bottom-up route (Lisman and Grace 2005). In the top-down route of the hippocampus-VTA loop, hippocampus indirectly projecting to the VTA, glutamate-releasing pyramidal neurons of the hippocampus (GLUergic neurons) innervate the median spiny neurons of the NAc (Lisman and Grace 2005). Neurons in NAc then send inhibitory GABAergic tone to the ventral Pallidum neurons, which in turn route inhibitory GABAergic tone onto VTA DA neurons (Frankle et al. 2006) (Lisman and Grace 2005). Alterations in the firing pattern of VTA DA neurons relays modulatory information back to the hippocampus, which defines one complete loop (Lisman and Grace 2005). Consequently, in the bottom-up route of this loop, VTA DA neurons directly innervate pyramidal neurons of the hippocampus and presumably mediate appetitive and motivational behaviors (Lisman and Grace 2005). Nevertheless, the role of the loop as a whole on reward-related learning process remains unknown. We hypothesized that the hippocampus-VTA loop bottom-up pathway could be the route of information flow via which the positive reinforcement properties of psychostimulants are mediated, whereas the top-down pathway attenuates the positive reinforcement properties of psychostimulants potentially by ensuing circuit-dependent disruptions of place learning. Disruptions in the circuit would hypothetically result in aversive behaviors that are associated with the intake of psychostimulants. Here, we show that the bottom-up pathway of the hippocampus-VTA loop mediates positive place reinforcement learning whereas the top-down pathway attenuates place learning via cellular mechanism that involves NMDA receptors.

Material and Methods

Subjects

Male Sprague-Dawley rats (325–349 g body weight upon arrival, Harlan Laboratories; $N = 80$) were housed two per cage

until surgery. Immediately after surgery and throughout the end of the experiments, the rats were kept individually. Their home cage room was maintained at constant temperature, 12-h light/dark cycle with food and water provided ad libitum. Prior to the start of any experiment, the rats were handled and acclimatized to a separate behavioral room by keeping them in the behavioral room for 2 h per day, for five consecutive days. All experimental protocols were approved in advance by the Institutional Animal Care and Use Committee and were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Surgeries and postoperative care

Before the start of all surgical procedures, Isoflurane gas anesthesia (Leica Microsystems Inc., Buffalo Grove, IL) was administered at full concentration (5%) and oxygen was administered at a 1.5 flow rate for a 5–7 min period through a Plexiglas chamber. Subjects received injections of rimadyl as an analgesic (rimadyl, 5 mg/kg, s.c.; Pfizer Animal Health, New York, NY) and baytril as an antibiotic (baytril, 2.5 mg/kg, i.p.; Bayer Animal Health, Pittsburgh, PA). Prior to mounting the subjects on a stereotaxic apparatus, the experimenter clipped hairs from the surgical sites, washed areas of incision at least three times by alternating betadine scrub, ethanol, and sterile water, and finally with iodine solution. The rats were then placed in a stereotaxic apparatus and the skin above the skull was incised. One (for VTA only) or three (VTA, VHC, NAc) small burr holes (3-mm diameter) were drilled above the skull for cannulae placement. Three sterile plastic guide cannula each containing a sterile stainless steel dummy (CMA/Microdialysis, Acton, MA) were aimed at the right hemisphere of each brain area of interest as follows (dimensions in mm): (a) VHC: A/P -4.0 , M/L $+3.5$, D/V -6.0 ; (b) the VTA: A/P -5.2 , M/L $+0.8$, D/V -6 ; and (c) the NAc: A/P $+1.5$, M/L $+2.5$, D/V -6.0 (Paxinos and Watson, 1998). The guides were slowly lowered to target nuclei via the holes and finally secured to the skull using bone screws and dental acrylic cement. During the postoperative care and treatment, rats were given once daily injections of rimadyl as analgesic (5 mg/kg s.c.) and baytril as antibiotics (2.5 mg/kg i.p.) for seven consecutive days. Occasionally and when necessary, baytril solution was added to water bottles (0.36 mL of the injectable form in 250-mL water bottles) for postoperative symptoms including loss of appetite, hair discoloration, or dehydrated skin.

Behavioral apparatus

Place conditioning: The apparatus is made of Plexiglas and was partitioned into three chambers (Fig. S1); black (left), gray (center), and white (right). The black and white chambers were equal in size ($26 \times 22 \times 33$ cm, each), while the

central chamber was smaller (18 × 22 × 33 cm) (Ricoy and Martinez 2009). The entire CPP apparatus was purchased from San Diego Instruments (San Diego Instruments, San Diego, CA) and had a Photo Activity System and software (PAS) that detects locomotion beam breaks and time spent in each chamber. The black and white chambers each had six photo beam sensors whereas the neutral central chamber had four. Previous studies from our laboratory (Ricoy and Martinez 2009) and our current preliminary data showed that rats show place bias for one of the two ends of the CPP apparatus, with most of the rats significantly preferring the black compared to the white chamber.

Behavioral assay

Intracranial conditioned place preference (IC-CPP): IC-CPP was used as a behavioral model of place reinforcement learning, modified from Ricoy and Martinez, 2009 (Ricoy and Martinez 2009). IC-CPP consisted of three main phases: (a) preconditioning, up to 3–4 days of familiarization for establishing the baseline place preference (30 min/session/day); (b) two consecutive days of conditioning (15 min/day) each followed by immediate testing (30 min/session/day); and (c) CPP testing 24 h and/or 1 week following conditioning (Experimental design, Fig. 1).

Preconditioning phase

The training for habituation takes three to four consecutive days depending on how long it takes for the rats to fulfill the criteria for baseline preference. The working criteria to achieving baseline habituation were defined as follows:

- (1) Average time spent (30 min/session/day) in the black chamber (preferred) increases from day-to-day while that of the white (nonpreferred) decreases accordingly; this means, rats must show a trend of habituation,
- (2) Average time spent in the preferred chamber should be significantly greater than that of the nonpreferred chamber,
- (3) The data collected 24 h before the commencement of IC-CPP experimental procedures were used as the baseline place preference, which was the reference point to compare the effect of the reinforcer on natural place preference. The reinforcer was METH or METH combined with MK801.

Conditioning phase

Reverse microdialysis application of METH (15 min/conditioning session) was used to apply the drug (Fig. 1). The reverse dialysis technique of IC-METH-CPP was previously used in our laboratory for similar behavioral studies (Ricoy and Martinez 2009). During conditioning, the infusion pump was turned ON for applying the drug via tiny diameter tubes (CMA microdialysis, FEP-tubing, volume 1.2 $\mu\text{L}/100\text{ mm}$) at the concentration of 10 $\mu\text{g}/\mu\text{L}$ and rate of 2.0 $\mu\text{L}/\text{min}$ for a total duration of 15 min. To be consis-

tent with our previous report (Ricoy and Martinez 2009), the concentration used was kept constant throughout (300 $\mu\text{g}/\text{session}$) but we did not measure the dose due to technical difficulties. During the 15-min conditioning, the rats were restrained within the nonpreferred chambers (against their baseline preference), whereas the Ringer's subjects (controls) were restrained within the preferred chambers. The same volume, rate of flow, and duration of conditioning were used for Ringer's groups (Ring). When the 15-min conditioning was completed, the microdialysis probes were carefully taken out and the guides were plugged with dummies, rats were then removed from the conditioning chambers, gently placed in the neutral chambers, and signal for START session sent from the computer, immediately. We did not assess all possible order of conditioning the circuit of interest (3!; six possible orders). Rather, we focused on changing the order of the VTA for the VHC and vice versa, and then maintained the order of conditioning the NAc constant (third order).

Testing phase

Behavioral data collection for CPP is often days, weeks, and even months apart from the actual conditioning date. However, in the current paradigm, we also tried to incorporate those relatively immediate behavioral alterations during the early stages (onsets) of place preference learning, before the drug is metabolically degraded from the brain. Thus, the IC-CPP training and testing protocol used in this project was slightly different from methods used by Ricoy and Martinez, 2009 (Ricoy and Martinez 2009). The testing module started immediately after establishing the baseline place preference as defined above in "Preconditioning phase" and following the conditioning as described in "Conditioning phase". Testing sessions were as follows: Test #1, immediately after conditioning #1; Test #2, immediately after conditioning #2; Test #3, (no treatment), 24 h following conditioning; and, Test #4, (no treatment), a week following conditioning; Note that to test the role of each nucleus within the hippocampus-VTA loop, testing was repeated according to the module three times. This means one module per brain area tested (experimental design, Fig. 1).

Drugs

Ringer's vehicle solution, Ring (Baxter, Deerfield, IL) was used to mimic artificial cerebrospinal fluid. Ringer's was composed of (mg/mL): 6 NaCl, 3.1 Sodium lactate, 0.3 KCl, and 0.2 CaCl₂. Dextromethamphetamine hydrochloride (METH, Sigma Chemical Co., St. Louis, MO) was dissolved in Ringer's at the concentration of 10 $\mu\text{g}/\mu\text{L}$. METH was prepared daily. The NMDA receptor noncompetitive antagonist MK-801 (Sigma Chemical Co., St. Louis, MO) was dissolved in

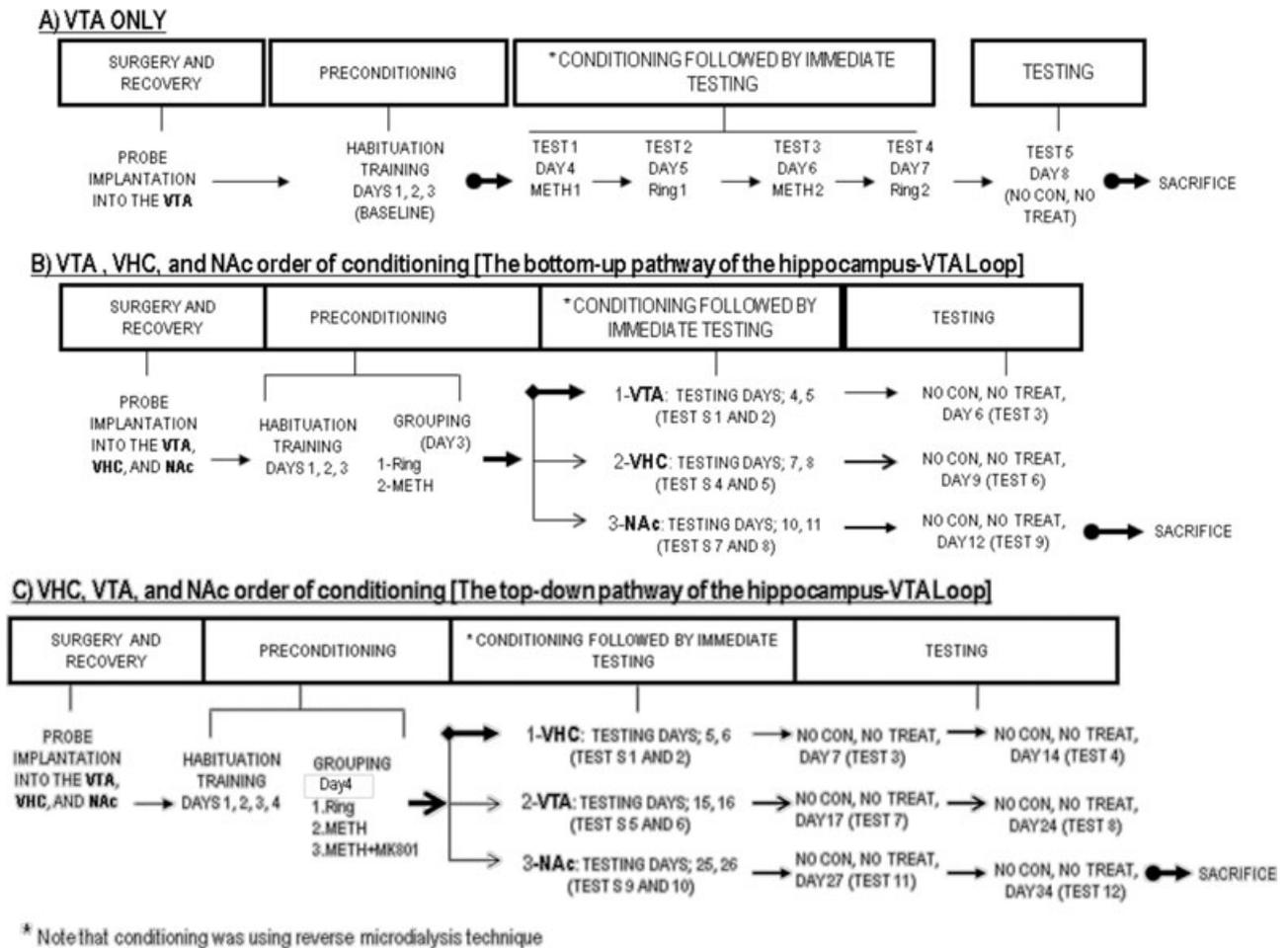


Figure 1. Timeline and experimental design: (A to C) flow chart of experimental design. All experimental rats underwent stereotaxic surgeries for intracranial probe implantations into the desired brain areas, all rats allowed to recover, and all were trained for baseline habituation: (A) following baseline training, the rats underwent four consecutive days of conditioning each followed by immediate testing by alternating METH with Ringer's (intra-VTA only). (B and C) Assessment on the hippocampus-VTA loop: rats were divided into groups immediately after establishing the baseline place preference, (B) VTA, then VHC, finally NAc (bottom-up) order of conditioning: two conditioning days each followed by immediate testing sessions (test 1 and test 2) and one CPP testing 24 h following conditioning (test 3). (C) VHC, then VTA, finally NAc (top-down) order of conditioning: two conditioning days each followed by immediate testing sessions (test 1 and test 2) and two CPP testing 24 h (test 3) and 1 week (test 4) following conditioning, performed for each brain area of interest. Also, note that the nucleus accumbens (NAc) was included in both the bottom-up and top-down pathways, in the same order (third order), to investigate the effect of order of exposure to METH on the expression of METH-induced CPP: VHC, ventral hippocampus; VTA, ventral tegmental area; NoCon, no conditioning; No Treat, no treatment but CPP testing performed.

freshly prepared METH solution (1:1) at a concentration of 0.1 mM.

Histology

After the behavioral experiments were completed rats were euthanized using isoflurane gas anesthesia. Brains were then removed immediately and preserved in methylbutane solution and stored at -80° in a freezer (no perfusion). Coronal sections (100- μ m thickness) were mounted onto gelatin-coated slides and subsequently stained with cresyl violet for verification of cannulae tip placements (Fig. S2).

Statistical analysis

Raw data were analyzed using one-way analysis of variance (ANOVA) for repeated measurements to determine if the groups differ in preferences before, during, and after conditioning. When significant main effects were detected, a Fisher's LSD post hoc tests were used for preplanned pairwise comparisons at $\alpha = 0.05$. To be consistent with our previous report (Ricoy and Martinez 2009), we used time deviation from baseline to represent either place preference (positive values) or place aversion (negative values). Thus, compared

to the baseline, a significant increase in time deviation in favor of the drug-paired chambers was interpreted as positive CPP whereas a significant decrease in time deviation was interpreted as place aversion.

Results

Intra-VTA reverse microdialysis application of METH produces positive place reinforcement learning

The VTA is functionally involved in the early stages (acquisition) of reinforcement learning. We previously reported that intra-VTA application of METH produced positive CPP in subjects that self administered cocaine, intravenously ($n = 5$; Society for Neuroscience conference, Keleta et al., 2009). In an extension and replication of the data, we used another set of rats that had no previous cocaine exposure (cocaine naïve). We hypothesized that METH applied intra-VTA would produce positive CPP learning even in subjects that had no previous cocaine experience. We used an “all-in-all-out” method of conditioning and testing by alternating METH with Ringer’s every other day, for four consecutive days. On the last day of testing, all subjects were tested for CPP without any intra-VTA treatment (Fig. 1A). Based on criteria described in “Behavioral Assay,” the rats satisfied the requirement for baseline place preference (Fig. 2A). There was a significant interaction between treatment (METH, Ringer’s) and Test (test 1, test 2) ($F [4, 48] = 5.03, P < 0.005, n = 13$). In agreement with our previous findings, METH reverse dialyzed into the VTA produced a markedly significant positive CPP. Compared to the baseline, a single (first time; METH1) intra-VTA conditioning session significantly increased the time deviation values in favor of the drug-paired chambers ($P < 0.05$). The METH1 groups, but not the Ringer’s1 groups, showed a positive CPP toward the METH-paired chambers compared to the baseline condition ($P < 0.05$) and compared to the Ringer’s group ($P < 0.05$). On repeated exposure with METH (METH2), however, the place conditioning effect was not different from that of Ringer’s and that of the baseline (Fig. 2 B–D). Following the second conditioning session with Ringer’s (Ringer’s2), rats showed positive bias toward the drug-paired chambers compared to that of the nonpaired chambers ($P < 0.05$), suggesting that METH-seeking behavior or withdrawal induced METH-seeking behavior. This later observation furthermore suggests that the novelty component of the reinforcer diminishes with repeated exposure and that the VTA primarily involves the detection of the novelty component of the METH. When tested 24 h following conditioning, without intra-VTA treatment, greater amount of time deviation values were found in the METH-paired chambers compared to the Ringer’s-paired chambers ($P < 0.05$). Overall, the observed deviation in place preference in favor of the drug-paired chambers can suggest that environ-

mental cues (the conditioned stimulus, CS+) paired with the METH (unconditioned stimulus, US+) produce stronger effect on drug-seeking behavior even in the absence of the reinforcer (US–).

The bottom-up pathway of the hippocampus-VTA loop mediates positive place reinforcement learning

As described in the Introduction, the hippocampus-VTA loop is implicated in the detection of novel stimulus entry into the CNS (Lisman and Grace 2005). The bottom-up pathway of this loop includes DAergic projections to the hippocampus and other cortical brain areas (Lisman and Grace 2005). If the novelty detection hypothesis (Lisman and Grace 2005) works then conditioning upstream of the comparator region should not affect novelty detection and hence should maintain the place reinforcing effects of METH. Consistent with this hypothesis, our finding in “Intra-VTA reverse microdialysis application of METH produces positive place reinforcement learning” suggests that stimulating the VTA produced positive CPP potentially because stimulation did not perturb the novelty comparator region of the hippocampus and hence the memory of the appetitive properties of METH remained intact. We therefore hypothesized that conditioning the bottom-up pathway of the hippocampus-VTA loop produces positive reinforcement learning following conditioning each of the three brain areas of interest within this loop. To do so, we conditioned another batch of rats in the order of VTA first followed by the VHC, and finally the NAc (refer Fig. 1B for experimental design). The following three successive experiments (“METH produced positive place learning following conditioning the VTA,” “In rats previously trained with intra-VTA-METH CPP, intra-VHC-METH produced positive place reinforcement learning 24 h following conditioning,” and “In rats previously trained with intra-VTA-METH followed by intra-VHC METH, intra-NAc-METH also produced an augmented positive place reinforcement learning 24 h following conditioning”) assessed the role of each of the three brain areas in METH-induced CPP learning.

METH produced positive place learning following conditioning the VTA

Based on criteria described in “Behavioral Assay,” the rats satisfied the requirement for baseline place preference (Fig. 3A). The rats in each group underwent intra-VTA CPP followed by testing. There was a significant interaction between treatments (Base [$n = 11$], Ringer’s [$n = 7$], METH [$n = 10$]) and Test (test 1, test 2, test 3) ($F [6, 46] = 8.74, P < 0.001$). In agreement with the above experiments in part I, the first intra-VTA conditioning session with METH, but not with Ringer’s, increased the time deviation values ($P < 0.001$).

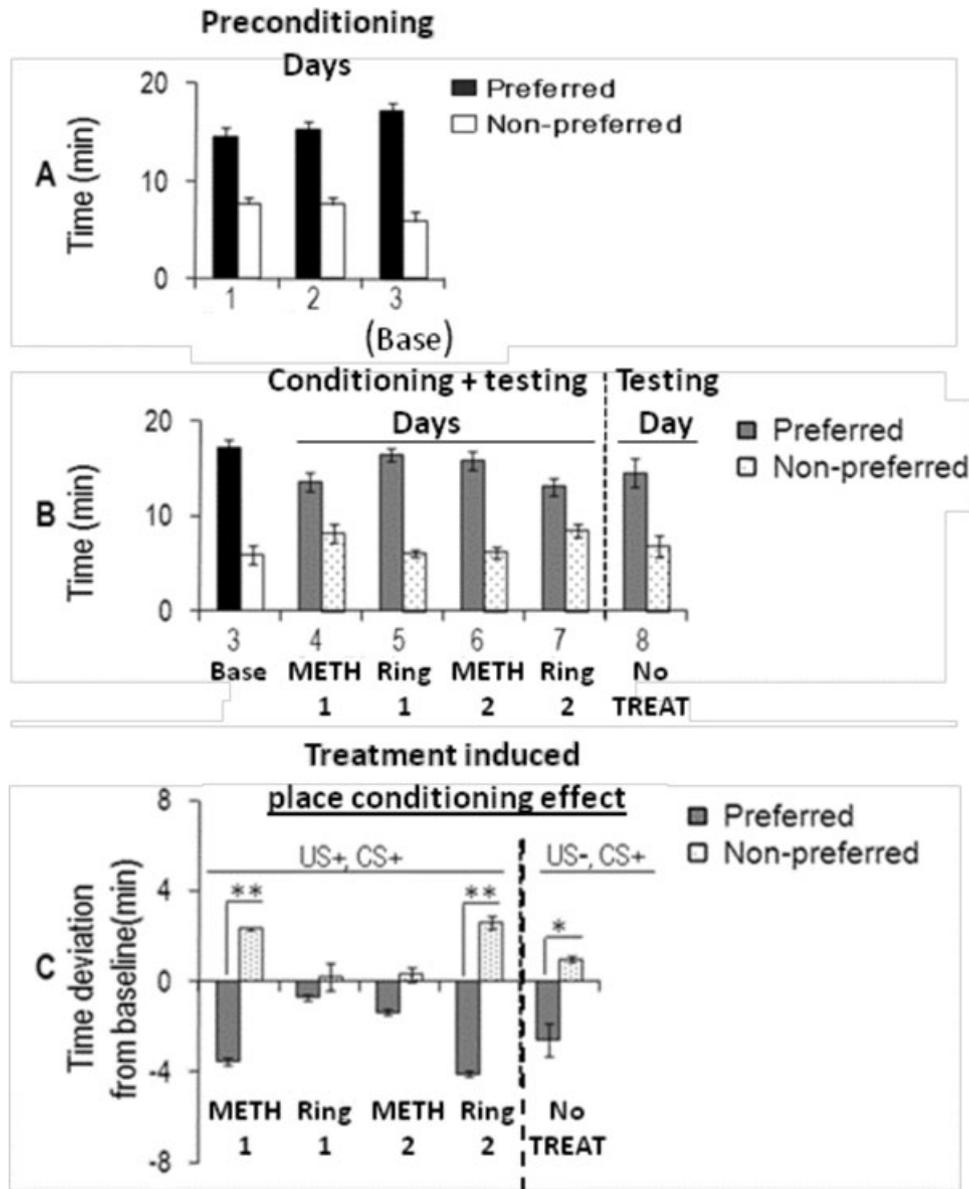


Figure 2. Intra-VTA METH induces positive place reinforcement learning. (A) Baseline place preference as defined by the amount of time per session prior to the commencement of IC-CPP. The rats were allowed to freely access the entire CPP runway to establish the baseline preference (base, day 3). Three consecutive days into the training (30 min/session/day), rats showed an increasing trend of place preference toward the black chambers (preferred) and a declining trend for the white chambers (nonpreferred). (B) The total time spent in the Ringer's-paired (gray bars) and METH-paired chambers (white dotted bars) following conditioning with either Ringer's or METH, and (C) the time deviation from baseline preference, in the Ringer's-paired chambers (gray bars) and in the METH-paired chambers (white dotted bars) normalized to the baseline. CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as standard error of the mean \pm SEM, $*0.005 < P < 0.05$.

The place conditioning effects of METH were also significantly greater than the baseline condition ($P < 0.05$). Additionally a positive increase in the time deviation from baseline was observed in the METH-paired chambers compared to the Ringer's-paired chambers ($P < 0.001$) (Fig. 3B–D). When tested 24 h following conditioning, without intra-

VTA treatment, METH-treated rats, but not Ringer's rats, showed increased time deviation values toward the METH-paired chambers ($P < 0.005$). The place reinforcing effects of METH was also greater than the baseline condition ($P < 0.05$) (Fig. 3E). Consistent with the previous experiment our findings showed that one intra-VTA application of METH was

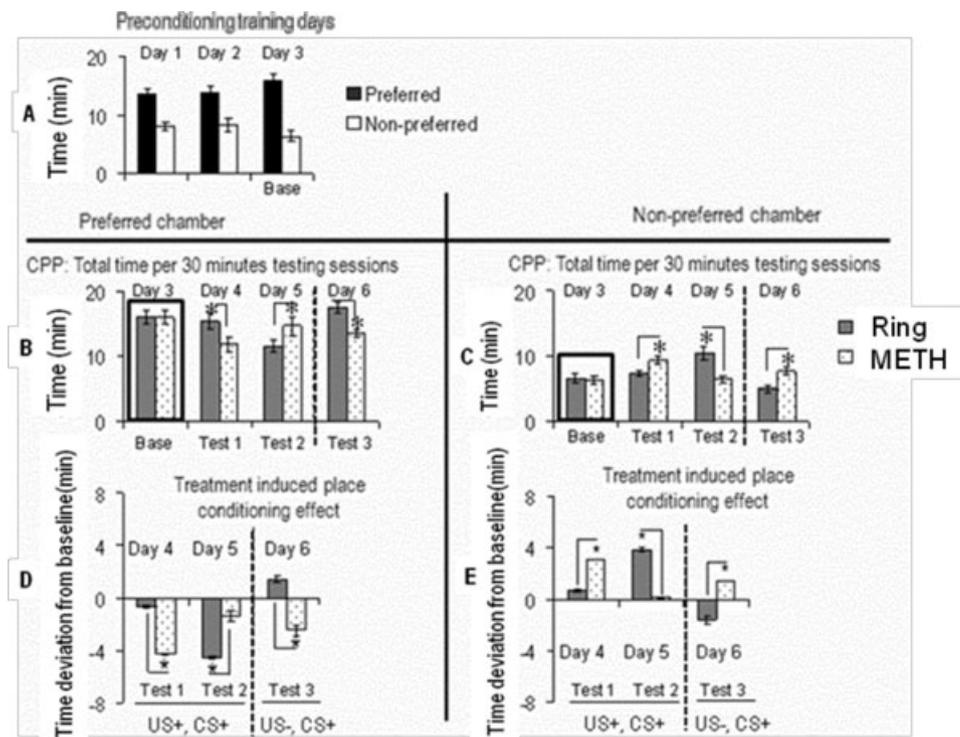


Figure 3. The bottom-up pathway of the hippocampus-VTA loop mediates positive place reinforcement learning following conditioning the VTA. (A) Baseline place preference is defined by the amount of time per session prior to the commencement of IC-CPP. Rats were allowed to freely access the entire CPP arena to establish the baseline preference. In three consecutive days of habituation, rats showed an increasing trend of place preference toward the black chambers and a declining trend for the white (B and C) the total amount of time spent (30 min/session/day); (B) in the Ringer's-paired, and (C) in the METH-paired chambers following conditioning with either Ringer's (gray bars) or METH (white dotted bars). (D and E) Time deviation from baseline preference: (D) in the Ringer's-paired chambers, and (E) in the METH-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, $*0.001 < P < 0.05$.

sufficient to produce positive CPP effects of METH expressed at least 24 h following conditioning.

In rats previously trained with intra-VTA-METH CPP, intra-VHC-METH produced positive place reinforcement learning 24 h following conditioning

After we finished our assessments on intra-VTA-METH-induced CPP learning, the same groups of rats from "METH produced positive place learning following conditioning the VTA" were conditioned with either METH or Ringer's intra-VHC, for the first time (refer Fig. 1B). There was a significant interaction between treatments (Base [$n = 11$], Ringer's [$n = 9$], METH [$n = 10$]) and test (Test 4, Test 5, Test 6) ($F [6, 49] = 3.39, P < 0.01$). Following the first-time intra-VHC exposure, the two groups did not statistically differ from one another, but both groups showed significant positive CPP toward the drug-paired chambers compared to the baseline condition ($P < 0.005$). The time deviation for the METH-

paired chambers following the second conditioning session was significantly reduced to a negative value below baseline ($P < 0.005$), however, there were no significant differences between METH-paired and Ringer's-paired groups on time deviation from the baseline condition ($P = 0.67$). To our surprise, 24 h following conditioning, METH rats, but not Ringer's rats, spent a significantly greater amount of time in METH-paired chambers compared to both the Ringer's group ($P < 0.05$) and the baseline condition ($P < 0.05$) (Fig. 4 B–D). In addition, METH groups spent a significantly more time in METH-paired chambers compared to Ringer's-paired chambers ($P < 0.01$).

In rats previously trained with intra-VTA-METH followed by intra-VHC METH, intra-NAC-METH also produced an augmented positive place reinforcement learning 24 h following conditioning

The NAc is highly implicated in the expression (or maintenance phase) of addictive behaviors associated with

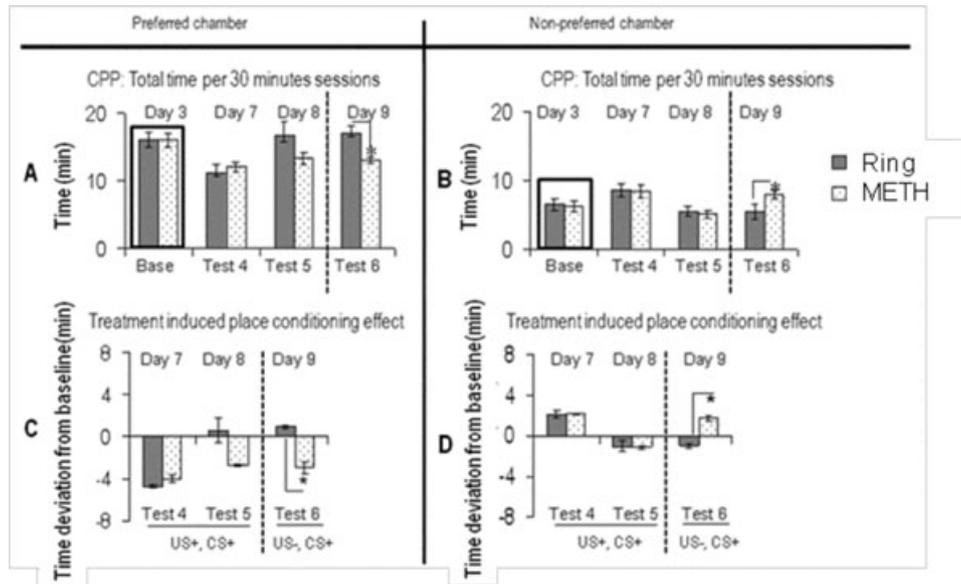


Figure 4. The bottom-up pathway of the hippocampus-VTA loop mediates positive place reinforcement learning following conditioning the VHC. (A and B) Total amount of time spent (30 min/session/day); (A) in the Ringer's-paired, and (B) in the METH-paired chambers following conditioning with either Ringer's (gray bars) or METH (white dotted bars). (C and D) Time deviation from baseline preference: (C) in the Ringer's-paired chambers, and (D) in the METH-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, * $P < 0.05$.

substances of abuse including METH (Rodriguez et al. 2008). Thus, to see the effect on the maintenance of IC-METH-CPP learning, we continued the experiment by finally conditioning the NAc. Therefore, the same rats from "In rats previously trained with intra-VTA-METH CPP, intra-VHC-METH produced positive place reinforcement learning 24 h following conditioning" were conditioned and tested with either METH or Ringer's intra-NAc, for the first time (Fig. 1B). There was a statistically significant interaction between treatments (Base [$n = 11$], Ringer's [$n = 9$], METH [$n = 10$]) and Test (test 7, test 8, test 9) ($F [6, 43] = 2.75$; $P < 0.05$). METH rats spent significantly more time in METH-paired chambers compared to Ringer's-paired chambers ($P < 0.001$). Moreover, compared to the Ringer's group and the baseline, the METH group showed a significant increase in time deviation toward METH-paired chambers 24 h following conditioning ($P < 0.001$) (Fig. 5 B–D). As expected, the observation following conditioning the NAc was robust.

The top-down pathway of the hippocampus-VTA loop down regulates place reinforcement learning

In the top-down order of conditioning (the VHC first, followed by the VTA, and finally the NAc), the hippocampus gets

METH treatment earlier than the VTA (experimental design; compare Fig. 1B and 1C). Here, the hypothesis is that disruption in the order of conditioning of the hippocampus-VTA loop would also disrupt METH-induced IC-CPP learning. Consequently, we investigated the role of the hippocampus-VTA top-down pathway in IC-METH-CPP learning. Moreover, as briefed in the Introduction section, one likely cellular mechanism by which psychostimulants affect long-term plasticity is by involving NMDA receptors. Hence, we in addition addressed the role of NMDA receptors on METH-induced CPP learning by combining METH with the NMDA receptor noncompetitive antagonist MK801. The long-term effect of METH-induced CPP was also addressed by testing CPP 1 week following conditioning each brain area of interest (Fig. 1C). Therefore, the next three successive experiments ("Intra-VHC-METH diminished place reinforcement learning which was reversed by NMDA receptor blockade," "In rats that were previously conditioned with intra-VHC-METH, intra-VTAMETH produced place aversion which was reversed by NMDA receptor blockade," and "In rats that were previously conditioned with intra-VHC followed by Intra-VTA-METH, intra-NAc-METH further produced place aversion which was reversed by NMDA receptor blockade.") assessed if conditioning the circuit in the order of VHC first followed by VTA and finally NAc produces CPP learning or not.

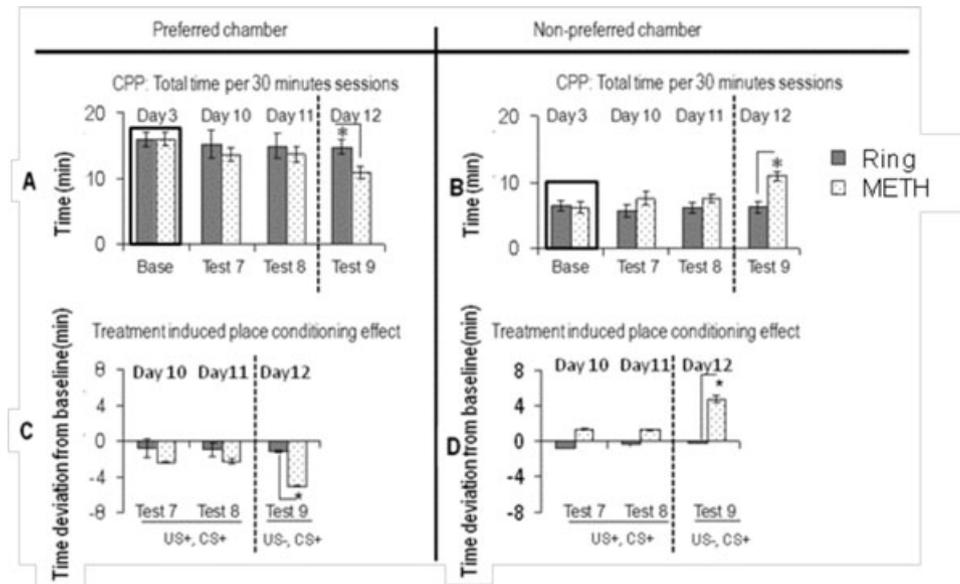


Figure 5. The bottom-up pathway of the hippocampus-VTA loop mediates positive place reinforcement learning following conditioning the NAc. (A and B) The total amount of time spent (30 min/session/day) (A) in the Ringer's-paired, and (B) in the METH-paired chambers following conditioning with either Ringer's (gray bars) or METH (white dotted bars). (C and D) The time deviation from baseline preference: (C) in the Ringer's-paired chambers, and (D) in the METH-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, * $P < 0.001$.

Intra-VHC-METH diminished place reinforcement learning which was reversed by NMDA receptor blockade

Based on criteria described in "Behavioral Assay", these rats (new batch of rats) satisfied the requirement for baseline place preference (Fig. 6A). The rats underwent intra-VHC conditioning against their initial place preference followed by immediate IC-CPP testing (30 min/session/day), while the controls (Ringer's group, $n = 6$) were conditioned within their preferred chambers. After two consecutive days of treatment (test 1 and test 2), no significant interaction between treatments was detected ($F [5, 22] = 0.43, P > 0.05$), however 24 h following conditioning, the METH+MK801 group ($n = 5$), but not the Ringer's ($n = 6$) and or the METH-treated groups ($n = 5$), showed positive CPP toward the drug-paired chambers ($P < 0.001$). In addition, METH+MK801 rats showed a statistically greater increase in time deviation toward the drug-paired chambers compared to the controls ($P < 0.001$). One week following conditioning, only the previously METH-treated rats showed positive bias toward Ringer's-paired chambers compared to both the METH+MK801 groups ($P < 0.001$) and the Ringer's groups ($P < 0.05$) (Fig. 6D and 6E). Our observation overall indicated that blocking the NMDA receptors reversed the diminished place learning following intra-VHC-METH. This attenuation in place learning could therefore be an

NMDA receptor activation-mediated process. The observation could also be METH-induced place aversion. Alternatively, because the initial place preference was negative relative to the positively conditioned side of the apparatus, the finding could be a block of CPP and that the data may reflect a block of learning rather than an aversion.

In rats that were previously conditioned with intra-VHC-METH, intra-VTA-METH produced place aversion which was reversed by NMDA receptor blockade

After we completed conditioning the VHC, we continued conditioning the same rats from "Intra-VHC-METH diminished place reinforcement learning which was reversed by NMDA receptor blockade", this time by applying METH intra-VTA (Fig. 1C). Twenty-four hours postintra-VTA conditioning, the METH-treated rats, but not the other two groups, showed increased time deviation toward the Ringer's-paired chambers (potentially aversion), ($P < 0.05$), whereas METH+MK801 group, but not the other two groups, showed a significant place preference toward drug-paired chambers ($P < 0.05$) a week following conditioning. Moreover, compared to METH+MK801 groups, the METH-treated rats showed significantly decreased time deviations away from the METH-paired chambers ($P < 0.01$) (Fig. 7A–D). The

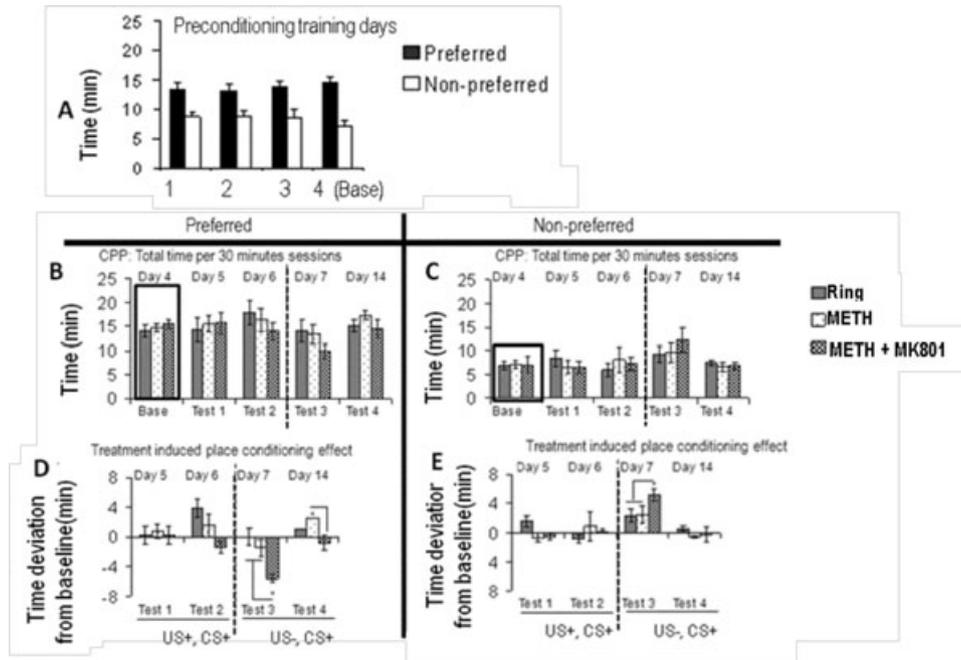


Figure 6. The top-down pathway of the hippocampus-VTA loop attenuates place preference learning following conditioning the VHC, which was reversed by inhibiting NMDA receptors using MK801. (A) Baseline place preference as defined by the amount of time per session prior to the commencement of IC-CPP testing. The rats were allowed to freely access the entire CPP runway to establish the baseline following habituation. After four consecutive days the rats showed an increasing trend of preference toward the black chambers (preferred) but a declining trend for the white chambers (nonpreferred). (B and C) The total amount of time spent (30 min/session/day); (B) in the Ringer's-paired, and (C) in the drug-paired chambers following conditioning with either Ringer's (gray bars), METH (white dotted bars), or METH+MK801 (black dotted bars). (D and E) The time deviation from baseline preference: (D) in the Ringer's-paired chambers, and (E) in the drug-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, $*P < 0.05$.

combination group (METH+MK801) preferred drug-paired chambers compared to the other two groups suggests that the METH induced place aversion that is potentially due to the activation of NMDA receptors. Overall, unlike the effect of METH on the induction of positive CPP that we observed following conditioning the bottom-up pathway of the hippocampus-VTA loop, METH in the top-down order of conditioning had place aversion effects even postconditioning the VTA (Fig. 7C and 7D).

In rats that were previously conditioned with intra-VHC followed by intra-VTA-METH, intra-NAC-METH further produced place aversion which was reversed by NMDA receptor blockade.

The next two consecutive days, day 21 and day 22, we continued conditioning the same rats from "In rats that were previously conditioned with intra-VHC-METH, intra-VTAMETH produced place aversion which was reversed by NMDA receptor blockade," this time by applying METH intra-NAC (Fig. 1C). The data showed that compared to the baseline, METH group rats had statistically significant place

aversion on test 9 ($P < 0.05$) and on test 10 ($P < 0.05$). As speculated and unlike the observation from postconditioning the NAc in the bottom-up order, intra-NAC-METH in the top-down order maintained place aversion while MK801 antagonized these effects (Fig. 8A-D). Surprisingly, this place aversion was expressed only when the drug is present in the brain (US+, CS+). In the absence of the US (US-), all the three groups came back to the baseline place preference, because there was no statistical differences between groups 24 h following conditioning (test 11, $P > 0.05$) and a week (test 12, $P > 0.05$). The major observation following conditioning the descending pathway (top-down) of the hippocampus-VTA loop is that conditioning the VHC earlier than the VTA produced aversive behaviors or at least attenuated the positive CPP induction properties of METH that we observed following conditioning the ascending pathway (bottom-up). Counter intuitively, the aversion response was more pronounced following conditioning the VTA and the NAc; however, this aversion response could be rescued by the blockade of NMDA receptors using the noncompetitive antagonist MK801.

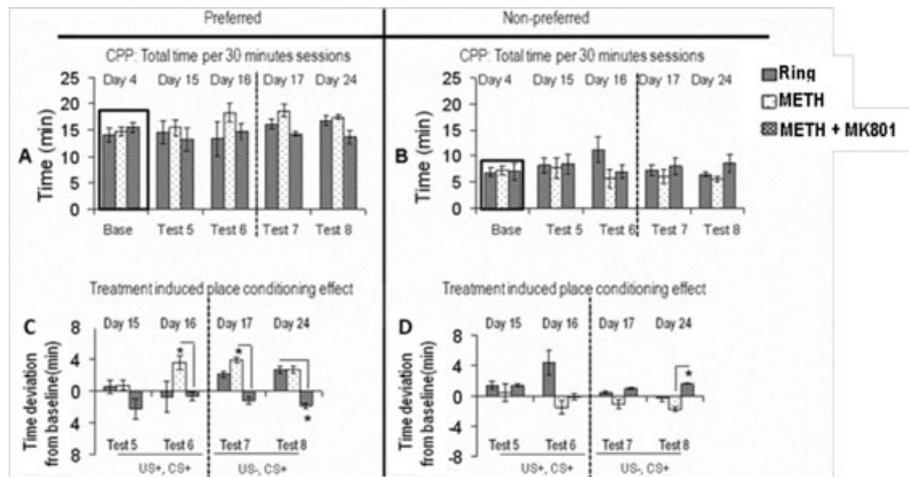


Figure 7. The top-down pathway of the hippocampus-VTA loop mediates place aversion following conditioning the VTA, which was reversed by inhibiting NMDA receptors using MK801. (A and B) the total amount of time (30 min/session/day) spent, (A) in the Ringer's-paired and, (B) in the drug-paired chamber following conditioning rats with either Ringer's (gray bars), METH (white dotted bars), or METH+MK801 (dark dotted bars). (C and D) The time deviation from the baseline preference: (C) in the Ringer's-paired chambers, and (D) in the drug-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, $*P < 0.05$.

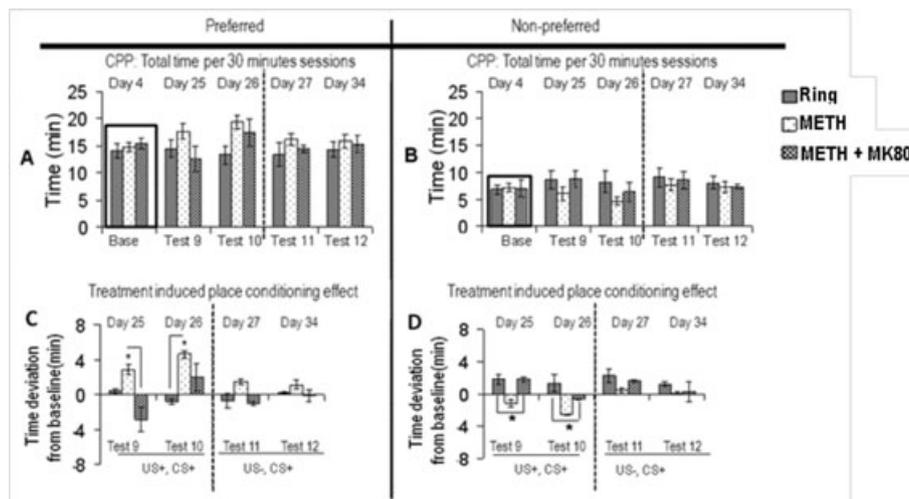


Figure 8. The top-down pathway of the hippocampus-VTA loop mediates place aversion following conditioning the NAc, which was reversed by inhibiting NMDA receptors using MK801. (A and B) total amount of time (30 min/session/day) spent, (A) in the Ringer's-paired, and (B) in the drug-paired chambers following conditioning with either Ringer's (gray bars), METH (white dotted bars), or METH+ MK801 (dark dotted bars). (C and D) time deviation from baseline preference: (C) in the Ringer's-paired chambers, and (D) in the drug-paired chambers. Abbreviations: CS+, conditioned stimulus present; US+, unconditioned stimulus present; US-, the unconditioned stimulus, METH, absent. Hatched vertical line separates treatment days (US+, CS+) from no treatment days (US-, CS+). Data are shown as \pm SEM, $*P < 0.05$.

Discussion

Cellular and electrophysiological mechanisms underlying place reinforcement learning within the hippocampus-VTA loop

The role of the NAc and mPFC in reward and motivation processing has been supported with pharmacological, molec-

ular, and electrophysiological lines of evidence. For instance, drugs of abuse (Ventura et al. 2001) and natural rewards (Gold et al. 1989) increase extracellular levels of DA in NAc. Rats self administer amphetamine into the NAc (Gilliss et al. 2002). In addition, electrophysiological investigations show that decreases in NAc GABAergic medium spiny neurons are implicated with reward-related behavioral responses. For

induces positive CPP. We also for the first time showed that the place reinforcement induction capacity of METH is dependent on the hippocampal-VTA loop. We observed that conditioning the bottom-up pathway of this loop, in the order of VTA first, followed by VHC, and finally NAc, produced positive place preference learning irrespective of where in the aforementioned three brain nuclei to which the drug was applied (Figs. 3, 4, and 5). In contrast, conditioning the top-down pathway of this loop, which involves earlier activation of the VHC using METH in the order of VHC, followed by VTA, and finally NAc, attenuated METH-induced positive CPP and thereby produced a place aversion (Figs. 6, 7, and 8). The aversive effects of METH in the top-down order of conditioning were attenuated by coadministration of METH with the NMDA receptor noncompetitive antagonist MK801 (in 1:1 ratio). This observation overall implies that there exists a Hebbian-type synaptic plasticity in which earlier activation of either of the two pathways in the hippocampus-VTA loop dominates and hence gradually produces an all-or-none plasticity; a plasticity of either positive place reinforcement learning (the bottom-up pathway) or plasticity of place aversion (the top-down pathway).

There might be a dorsoventral distinction of the hippocampal formation with respect to reinforcement learning

Our laboratory previously reported that amphetamine reverse-dialyzed intra-NAc produced positive CPP (Rodriguez 2008). In addition, using intrahippocampal METH self administration and intrahippocampal METH CPP behavioral paradigms, our laboratory reported that METH microdialyzed into the dorsal hippocampus-induced positive reinforcement learning. Ricoy and Martinez, 2009 (Ricoy and Martinez 2009) further reported that the positive reinforcement capacities of METH treatment within the dorsal hippocampus was a D1-like receptor-mediated process because the D1 receptor antagonist Schering, SCH23390, coadministered with METH attenuated the reinforcing efficacy of METH (Ricoy and Martinez 2009). However, other research shows that there is a dorsoventral functional segregation of the hippocampal structure in which the dorsal portion performs primarily motor-related cognitive functions, whereas the ventral portion mediates affective behaviors and emotions (Fanselow and Dong 0000). Unlike the case for dorsal hippocampus, our current findings following conditioning the hippocampus-VTA loop was an “if... then...” condition. If the midbrain DA system (VTA) were chemically stimulated, in this case with METH, earlier than the VHC, then METH produces positive place reinforcement effect regardless of where in these three regions it is first applied. By contrast, if the VHC were chemically stimulated with METH earlier than the VTA, then METH produces place aversion

learning; or it fails to produce positive place reinforcement even following the application of the drug into the other two regions (VTA and NAc). This observation led us to the hypothesis that the ascending pathway in the hippocampus-VTA loop mediates positive reinforcement learning while the descending pathway mediates aversion related to the exposure of psychostimulants such as METH. Needless to say, dose response effect of METH applied into each nucleus of the hippocampus-VTA loop is worth investigating perhaps by operant conditioning behavioral techniques such as intravenous or intracranial drug self administration.

Concluding remark

In summary, compared to the top-down order of conditioning, the bottom-up order of conditioning the hippocampus-VTA loop produced METH-mediated place reinforcement learning. The top-down order of conditioning attenuated place learning or produced place aversion. Blocking the NMDA receptors reversed this effect, which is interesting in light with the electrophysiological findings by Lisman and Grace (Lisman and Grace 2005). Because addiction is a learning process, we propose that disruption of the learning circuitry also disrupts learning. We assume that interfering with the natural flow of neural information during the process of novel stimulus entry to the CNS results in aberrant learning. Thus, the reinforcing properties of psychostimulants including METH can be attenuated when one begins stimulating from the top-down pathway, which is anatomically located downstream of the comparator region of the hippocampus-VTA loop (Lisman and Grace 2005).

Addiction is a very complicated psychological disease that becomes more and more complicated as time progresses. We believe that research on addiction should try to address this disease at its earliest phase (acquisition) rather than trying to find the solution at its full blown (expression) phase. On the basis of our current findings, we suggest that future investigations should focus on neural and behavioral correlates of the hippocampus-VTA loop with due emphasis given on the acquisition phase of reinforcement learning. The findings from such research projects would help us develop some target-specific (e.g., receptor or receptor subunit specific) therapeutics for addiction-related health problems and any other psychological disorders that emanate as a result of exposing the brain to psychoactive drugs.

Acknowledgments

The research project was supported by the diversity program in neuroscience at NIMH, The Ewing Halsell Distinguished Chair under the American Psychological Program, APA T32 MH18882 and by the MBRS program at University of Texas at San Antonio, NIH GM07717.

References

- Adamec, R. 2001. Does long term potentiation in periaqueductal gray (PAG) mediate lasting changes in rodent anxiety-like behavior (ALB) produced by predator stress?—Effects of low frequency stimulation (LFS) of PAG on place preference and changes in ALB produced by predator stress. *Behav. Brain Res.* 120:111–135.
- Bannerman, D. M., B. Niewoehner, L. Lyon, C. Romberg, W. B. Schmitt, A. Taylor, D. J. Sanderson, J. Cottam, R. Sprengel, P. H. Seeburg, G. Köhr, and J. N. Rawlins. NMDA receptor subunit NR2A is required for rapidly acquired spatial working memory but not incremental spatial reference memory. *J. Neurosci.* 28:3623–3630.
- Berridge, K. C., I. L. Venier, and T. E. Robinson. 1989. Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: implications for arousal and anhedonia hypotheses of dopamine function. *Behav. Neurosci.* 103:36–45.
- Broadbent, N. J., L. R. Squire, and R. E. Clark. 2004. Spatial memory, recognition memory, and the hippocampus. *Proc. Natl. Acad. Sci. USA* 101:14515–14520.
- Brown, T. E., B. R. Lee, and B. A. Sorg. 2008. The NMDA antagonist MK-801 disrupts reconsolidation of a cocaine-associated memory for conditioned place preference but not for self-administration in rats. *Learn. Mem.* 15:857–865.
- Cannon, C. M., and R. D. Palmiter. 2003. Reward without dopamine. *J. Neurosci.* 23:10827–10831.
- Carter, D. A., and H. C. Fibiger. 1977. Ascending projections of presumed dopamine-containing neurons in the ventral tegmentum of the rat as demonstrated by horseradish peroxidase. *Neuroscience* 2:569–576.
- Fanselow, M. S., and H. W. Dong. 2010. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65:7–19.
- Floresco, S. B., C. L. Todd, and A. A. Grace. 2001. Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *J. Neurosci.* 21:4915–4922.
- Frankle, W. G., M. Laruelle, and S. N. Haber. 2006. Prefrontal cortical projections to the midbrain in primates: evidence for a sparse connection. *Neuropsychopharmacology* 31:1627–1636.
- Gasbarri, A., C. Verney, R. Innocenzi, E. Campana, and C. Pacitti. 1994. Mesolimbic dopaminergic neurons innervating the hippocampal formation in the rat: a combined retrograde tracing and immunohistochemical study. *Brain Res.* 668:71–79.
- Gasbarri, A., A. Sulli, and M. G. Packard. 1997. The dopaminergic mesencephalic projections to the hippocampal formation in the rat. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 21:1–22.
- Gilliss, B., C. J. Malanga, J. O. Pieper, and W. A. Carlezon, Jr. 2002. Cocaine and SKF-82958 potentiate brain stimulation reward in Swiss-Webster mice. *Psychopharmacology* 163:238–248.
- Gold, L. H., M. A. Geyer, and G. F. Koob. 1989. Neurochemical mechanisms involved in behavioral effects of amphetamines and related designer drugs. *NIDA Res. Monogr.* 94:101–126.
- Itzhak, Y. 2008. Role of the NMDA receptor and nitric oxide in memory reconsolidation of cocaine-induced conditioned place preference in mice. *Ann. NY Acad. Sci.* 1139:350–357.
- Jenkins, T. A., E. Amin, J. M. Pearce, M. W. Brown, and J. P. Aggleton. 2004. Novel spatial arrangements of familiar visual stimuli promote activity in the rat hippocampal formation but not the parahippocampal cortices: a c-fos expression study. *Neuroscience* 124:43–52.
- Kauer, J. A., and R. C. Malenka. 2007. Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8:844–858.
- Lee, I., M. R. Hunsaker, and R. P. Kesner. 2005. The role of hippocampal subregions in detecting spatial novelty. *Behav. Neurosci.* 119:145–153.
- Lisman, J. E., and A. A. Grace. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46:703–713.
- Peoples, L. L., and M. O. West. 1996. Phasic firing of single neurons in the rat nucleus accumbens correlated with the timing of intravenous cocaine self-administration. *J. Neurosci.* 16:3459–3473.
- Ricoy, U. M., and J. L. Martinez. 2009. Local hippocampal methamphetamine-induced reinforcement. *Front. Behav. Neurosci.* 3(47):1–10.
- Rodriguez, J. S., S. Y. Boctor, C. F. Phelix, and J. L. Martinez, Jr. 2008. Differences in performance between Sprague-Dawley and Fischer 344 rats in positive reinforcement tasks. *Pharmacol. Biochem. Behav.* 89:17–22.
- Rossato, J. I., L. R. Bevilacqua, I. Izquierdo, J. H. Medina, and M. Cammarota. 2009. Dopamine controls persistence of long-term memory storage. *Science* 325:1017–1020.
- Schultz, W. 1998. Predictive reward signals of dopamine neurons. *J. Neurophysiol.* 80(1):1–27.
- Ventura, R., S. Cabib, and S. Puglisi-Allegra. 2001. Opposite genotype-dependent mesocorticolimbic dopamine response to stress. *Neuroscience* 104:627–631.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Schematic representation of the three chambered conditioned place preference apparatus along with the Photo Activated Software (PAS).

Figure S2. Coronal section representation of unilateral cannulae tip placement.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.