

Negative Reward Signals from the Lateral Habenula to Dopamine Neurons Are Mediated by Rostromedial Tegmental Nucleus in Primates

Simon Hong,¹ Thomas C. Jhou,³ Mitchell Smith,¹ Kadharbatcha S. Saleem,² and Okihide Hikosaka¹

¹Laboratory of Sensorimotor Research, National Eye Institute and ²Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, and ³Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina 29425

Lateral habenula (LHb) neurons signal negative “reward-prediction errors” and inhibit midbrain dopamine (DA) neurons. Yet LHb neurons are largely glutamatergic, indicating that this inhibition may occur through an intermediate structure. Recent studies in rats have suggested a candidate for this role, the GABAergic rostromedial tegmental nucleus (RMTg), but this neural pathway has not yet been tested directly. We now show using electrophysiology and anatomic tracing that (1) the monkey has an inhibitory structure similar to the rat RMTg; (2) RMTg neurons receive excitatory input from the LHb, exhibit negative reward-prediction errors, and send axonal projections near DA soma; and (3) stimulating this structure inhibits DA neurons. Surprisingly, some RMTg neurons responded to reward cues earlier than the LHb, and carry “state-value” signals not found in DA neurons. Thus, our data suggest that the RMTg translates LHb reward-prediction errors (negative) into DA reward-prediction errors (positive), while transmitting additional motivational signals to non-DA networks.

Introduction

Neurons in the primate lateral habenula (LHb) are excited by visual stimuli that predict the absence of reward and are inhibited by stimuli that predict the presence of reward (Matsumoto and Hikosaka, 2007). These patterns are inverse to those found in dopamine (DA) neurons, consistent with findings that LHb stimulation strongly inhibits dopamine neurons in the substantia nigra pars compacta (SNc) and ventral tegmental area (VTA) (Lisoprawski et al., 1980; Christoph et al., 1986; Ji and Shepard, 2007; Matsumoto and Hikosaka, 2007). Recent studies using rats found that the direct projection from the LHb to DA neurons is glutamatergic (Omelchenko et al., 2009; Brinschwitz et al., 2010). Therefore the suppressive effect of the LHb on DA neurons needs to be disynaptic or multisynaptic. One possibility is that the suppressive effect is mediated by GABAergic interneurons within the VTA. However, the analysis by Omelchenko et al. (2009) suggests that only about 16% of LHb axons terminate within the VTA and that these do not preferentially contact local GABAergic neurons, suggesting that the suppressive effect may be mediated through another brain structure.

One possible mediator of this inhibitory action is a structure in the mesopontine area caudal to VTA, which we will refer to as the rostromedial tegmental nucleus (RMTg) (Jhou et al., 2009a) [also called the “caudal tail of VTA” (Kaufling et al., 2009) and “paramedian raphe nucleus” (Paxinos et al., 1999; Kim, 2009)]. This structure is a good candidate because of (1) its heavy innervation by LHb inputs (Herkenham and Nauta, 1979; Jhou et al., 2009a; Kaufling et al., 2009; Kim, 2009), (2) its prominent projection to DA-rich areas (Jhou et al., 2009a; Kaufling et al., 2009; Balcita-Pedicino et al., 2011), and (3) its GABAergic nature (Perrotti et al., 2005; Olson and Nestler, 2007; Jhou et al., 2009a,b; Kaufling et al., 2009; Balcita-Pedicino et al., 2011). Furthermore, the similarity of reward and punishment responses between the RMTg neurons in the rat (Jhou et al., 2009b) and LHb neurons in the monkey (e.g., Matsumoto and Hikosaka, 2007; Hong and Hikosaka, 2008a) strongly suggests that the RMTg relays LHb signals to DA neurons. However, the functional connectivity of this hypothesized circuit has not been tested directly.

Because prior studies of the RMTg were done in rats, we sought to determine whether the primate has a homologous structure, and if so, whether it is functionally connected to LHb and DA neurons. We also examined whether the RMTg carries information related to other structures to which it is connected. Our methods to achieve these goals were threefold: (1) find and characterize neurons that encode reward-related signals by having the monkey perform a reward-biased eye movement task, (2) determine whether the reward-related neurons receive inputs from the LHb and send outputs to the SNc using orthodromic and antidromic stimulations, and (3) reconstruct the locations of these neurons histologically. We also conducted a retrograde tracer study to delineate the boundaries of the primate RMTg and

Received March 18, 2011; revised May 23, 2011; accepted June 13, 2011.

Author contributions: S.H. and O.H. designed research; S.H., T.C.J., M.S., and O.H. performed research; S.H., T.C.J., M.S., K.S.S., and O.H. contributed unpublished reagents/analytic tools; S.H., T.C.J., and O.H. analyzed data; S.H., T.C.J., M.S., K.S.S., and O.H. wrote the paper.

This work was supported by the intramural research program of the National Eye Institute. We are grateful to E. S. Bromberg-Martin, H. Kim, Y. Tachibana, S. Yamamoto, and M. Yasuda for helpful comments, and D. Parker, B. Nagy, G. Tansey, J. W. McClurkin, A. M. Nichols, and T. W. Ruffner for technical assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to Simon Hong, Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, 49 Convent Drive, Bethesda, MD 20892. E-mail: hongy@nei.nih.gov.

DOI:10.1523/JNEUROSCI.1384-11.2011

Copyright © 2011 the authors 0270-6474/11/3111457-15\$15.00/0

compared it with the reconstructed locations of the RMTg neurons.

Materials and Methods

Two male rhesus monkeys (*Macaca mulatta*), B and C, were used as subjects in this study. All animal care and experimental procedures were approved by the National Eye Institute and Institutional Animal Care and Use Committee and complied with the Public Health Service Policy on the humane care and use of laboratory animals.

Behavioral task. Behavioral tasks were the same as the ones described previously (Hong and Hikosaka, 2008a). The monkey was seated in a primate chair. Visual stimuli were rear projected by a projector onto a frontoparallel screen 33 cm from the monkey's eyes. Eye movements were monitored using a scleral search coil system. The monkey was trained to perform the one-direction-rewarded (IDR) task (see Results). For the monkey C, colors (blue for rewarding target, red for nonrewarding target) were added to the visual target. A trial started when a small fixation spot appeared on the screen. After the monkey maintained fixation on the spot for 750–1250 ms, the fixation spot disappeared, and a peripheral target appeared on either the right or left side, 10° from the fixation spot. The monkey was required to make a saccade to the target within 750 ms. Correct and incorrect saccades were signaled by a tone and a beep 200 ms after the saccade, respectively. Within a block of 24 trials, saccades to one fixed direction were rewarded with 0.3 ml of apple juice, whereas saccades to the other direction were not rewarded. The position-reward contingency was reversed in the next block with no external instruction. Even in the unrewarded trials, the monkey had to make a correct saccade; otherwise, the same trial was repeated. In rewarded trials, a liquid reward was delivered, which started simultaneously with a tone stimulus.

Electrophysiology. One recording chamber was placed over the midline of the parietal cortex, tilted posteriorly by 40°, and aimed at the Lhb and the SNc; the other recording chamber was placed over the frontoparietal cortex, tilted laterally by 35°, and aimed at the RMTg. Single-unit recordings and electrical stimulations were performed using tungsten electrodes that were advanced by an oil-driven micromanipulator. The neural signal was amplified with a bandpass filter and sampled at 40 kHz. Single neurons were isolated on-line using a custom voltage–time window discrimination software. DA neurons were identified by their irregular firing, tonic baseline activity around five spikes per second, broad spike potential, and phasic excitation to free reward (Matsumoto and Hikosaka, 2009).

Orthodromic and antidromic activation and collision. For the stimulation of the Lhb, the position of the Lhb was mapped first by MRI. The electrophysiological features of the Lhb (Matsumoto and Hikosaka, 2007) were also used to locate the Lhb. After finding the Lhb, the IDR task was performed, and single-unit or multiunit activity of the Lhb was recorded. After finishing the recording, the Lhb electrode was connected to the stimulator (S88; Grass Technologies). To minimize the electrical artifact, we used a commercially available artifact remover (Artifact Zapper-1; Riverbend Instruments). For stimulation, we delivered biphasic negative–positive pulses with 0.2 ms per phase duration between the Lhb electrode and the guide tube. The stimulation current was 10–200 μ A. To examine the orthodromic RMTg activation by the Lhb electric stimulation, the Lhb was stimulated every \sim 1.0 s after finding a presumed RMTg neuron (we often recorded presumed RMTg neurons with the IDR task at the expected depth of the RMTg site before confirming orthodromic modulation).

On some occasions, we examined antidromic activations of the Lhb as well. This was done by stimulating the RMTg and looking for any sign of antidromic spikes at the Lhb site. When we detected a spike consistently occurring with a fixed poststimulation latency, we tried to isolate the spike from the background activity using a voltage–time window discrimination software (MEX, developed by Laboratory of Sensorimotor Research, National Eye Institute, NIH). To confirm the connectivity, a collision test was performed by stimulating the RMTg a few milliseconds after detecting a spontaneously occurring Lhb spike. If the stimulation-evoked Lhb spike disappeared after decreasing the time between the detection of an Lhb spike and the delivery of RMTg stimulation, the

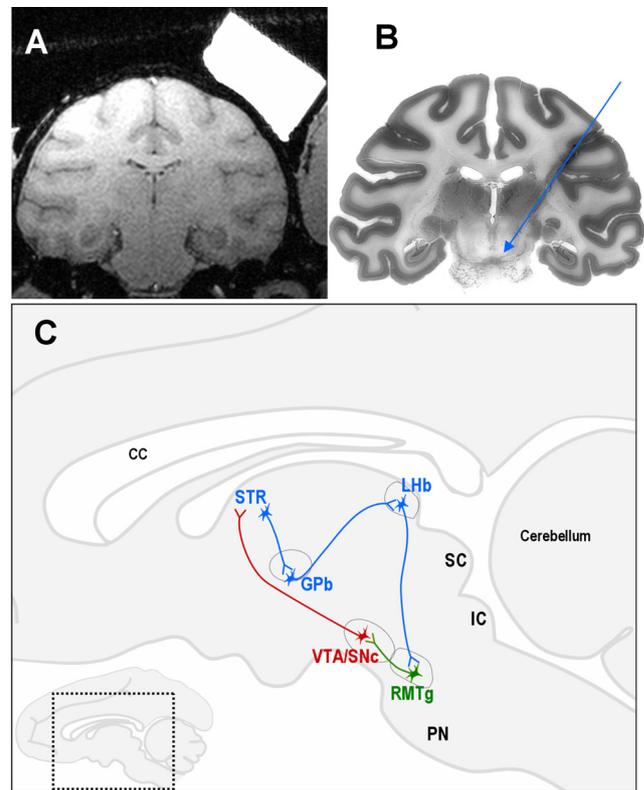


Figure 1. Neural recording configuration. **A** The MRI image of the brain section corresponding to the anatomical section in **B**. The white rectangular area on the upper right part of the brain is the recording chamber. The MRI chamber was filled with gadolinium to enhance its MRI image. The image is from monkey B. **B**, The schematic of our recording approach to the RMTg. The arrow indicates a representative path of the recording electrode. **C**, The gross anatomy of the circuit that we examined. STR, Striatum; GPb, border part of the globus pallidus; PN, pontine nuclei; SC, superior colliculus; IC, inferior colliculus; CC, corpus callosum.

spike was considered to be activated antidromically, provided that this collision latency was slightly longer than the antidromic latency by about 0.3 ms (absolute refractory period). Then, the Lhb neuron whose spike was activated antidromically was considered to project to the RMTg. We then recorded the activity of the Lhb neuron while the monkey was performing the IDR task.

After identifying the RMTg, electric stimulations were applied at the center of the structure while recording the activity of DA neurons in the SNc to examine orthodromic modulation of the DA neurons. Antidromic activation of RMTg neurons by the stimulation of the DA site was also performed occasionally.

Histological examination. For the histological reconstruction of the locations of the reward-related neurons, we relied on electrolytic lesions at recording sites and electrode tracks, both visualized in histological Nissl-stained sections. All sections except for one every five sections from the level of the subthalamic nucleus to the inferior colliculus (>200 sections, >10 mm) were Nissl stained, and all electrolytic lesions and electrode tracks were localized. The locations of recorded neurons were reconstructed using 3D coordinates read from the electrode manipulator with respect to the reconstructed positions of the electrolytic lesions.

Cholera toxin B subunit retrograde tracer study. The monkey B was used for the anatomical study. Using the orthodromic and antidromic activation/collision tests described above, the RMTg was identified. Upon the identification, electrolytic microlesions were made at the dorsal and ventral edges of the presumed RMTg. We then injected cholera toxin B subunit (CTB; 0.9%, 0.2 μ l) into the right SNc immediately after finding a typical reward-related DA neuron. The injection site was the estimated SNc site where we had observed bidirectional (orthodromic, antidromic)

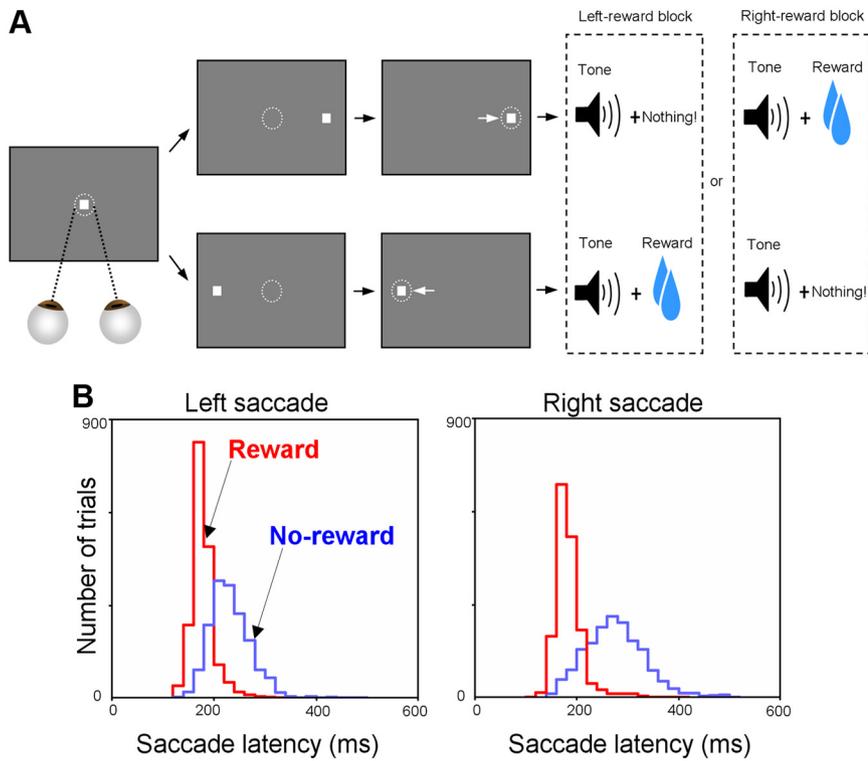


Figure 2. Behavioral task and animals' task performance. **A**, Sequence of events in the 1DR version of the visually guided saccade task. The monkey first fixated at the central spot (the dotted circle indicates the eye position). As the fixation point disappeared, a target appeared randomly on the right or left and the monkey, was required to make a saccade to it immediately. Correct saccades in one direction were followed by a tone and juice reward; saccades in the other direction followed by a tone alone. The rewarded direction was fixed in a block of 24 trials and was changed in the subsequent block. **B**, Distribution of saccade latencies in rewarded trials (in red) and in unrewarded trials (in blue) (data from monkey B). Saccades in the first trials after the changes in position-reward contingency have been excluded.

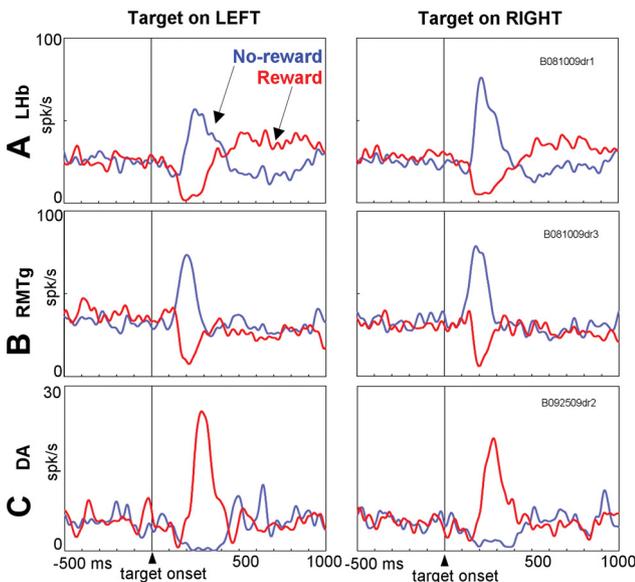


Figure 3. Responses of representative Lhb, RMTg, and DA neurons to target onset in the 1DR task. The averaged activity of each neuron, expressed as an SDF, is shown separately for the reward trials (red) and no-reward trials (blue) as the response to the onset of the left target (left) and the right target (right). The neurons were recorded from the left hemisphere.

stimulation effects between the SNc and the RMTg (see Results). After the conclusion of the experiment, each animal was deeply anesthetized with an overdose of pentobarbital sodium and perfused with 4% paraformaldehyde. The brain was blocked and equilibrated with 30% su-

crose. Frozen sections were cut every 50 μm in the plane parallel to the electrode penetration into the RMTg. Every fifth section was stained to visualize CTB. In brief, sections were incubated overnight at room temperature in goat anti-CTB primary (List Biological; 1:50,000 dilution in PBS with 0.25% Triton X-100) and rinsed three times for 2 min each in PBS, followed by incubation for 1 h in biotinylated donkey-anti-goat secondary antibody (Jackson ImmunoResearch; 1:1000). After six 2 min rinses in PBS, tissue was incubated for 1 h in avidin-biotin complex (Vector Laboratories), followed by three 2 min rinses in PBS and incubation in 0.05% diaminobenzidine with 0.01% hydrogen peroxide for 10–20 min. This final incubation revealed a brownish reaction product, after which sections were rinsed two times for 2 min each in PBS and then mounted and coverslipped. The rest of the sections were stained with thionin violet to identify lesion markings and electrode tracks.

Statistical analysis. We defined the post-target response as the average discharge rate during the 150–350 ms period after the target onset minus the background discharge rate measured during the 1000 ms period before the fixation point appeared. The reward response was defined as the average discharge rate during 150–350 ms after the onset of the tone stimulus (which was synchronized with reward onset if reward was present) minus the background discharge rate. We set the time windows such that they included major parts of the excitatory and inhibitory responses of Lhb, RMTg, and DA neurons.

Using one-way ANOVA and receiver operating characteristic (ROC) we classified RMTg neurons into three groups: (1) reward-positive type, if their reward modulation had positive values ($p < 0.01$, ANOVA; ROC, > 0.5); (2) negative type, if their reward modulation had negative values ($p < 0.01$, ANOVA; ROC, < 0.5); and (3) reward unmodulated type ($p > 0.01$, ANOVA). We further classified these reward-modulated neurons into “state-value” type and “change-of-value” type neurons. We considered the neurons having a significantly ($p < 0.05$, Wilcoxon signed rank test) lasting modulation after the appearance of the fixation stimulus as state-value neurons. The time window of this test was 350 to 750 ms after the appearance of the fixation stimulus. The 350 ms constraint was added because many neurons showed just brief phasic modulation shortly after the appearance of the fixation stimulus. Some of the neurons that met the above state-value criterion showed a reversal of modulations, which indicated a rebound activity following a phasic postfixation burst or suppression; we classified those neurons as change-of-value neurons.

We determined the latency of reward-dependent modulation for each of the three groups of neurons: negative RMTg neurons, positive RMTg neurons, and Lhb neurons. First, we quantified for each neuron, at each time point after target onset, how much its activity is different between rewarded trials and unrewarded trials. For this purpose we computed a spike density function (SDF) for each trial. Based on the trial-by-trial SDFs, we computed an ROC value at every 1 ms bin, starting from 1000 ms before target onset until 1000 ms after target onset. Using the two-tailed permutation test, we determined whether the ROC value comparing the rewarded and unrewarded trials was significantly separated from the ROC value based on the shuffled data ($p < 0.01$, with 1000 permutations). If the significant difference held true for 25 consecutive time bins (25 ms), we judged that the neuron showed significant reward-dependent modulation during the 25 ms period. This method efficiently eliminated occasional blips that reached the significance level (on average, 1% of the examined period is expected to be significant by defini-

tion). Then, for each group of neurons, we counted the number of neurons at each time bin that showed reward-dependent modulation. The latency of the reward-dependent modulation for each group of neurons was determined at the time point when the number of neurons that showed the reward-dependent modulation significantly exceeded the control variation level (an upper 1% SD level based on the data during the 1000 ms pretarget period) for at least 25 consecutive time bins.

To determine the latency of the orthodromic response of RMTg and DA neurons in response to the electrical stimulation of the LHB and the RMTg, respectively, we used the Poisson distribution test and the Wilcoxon signed rank test for excitation and inhibition, respectively. For Poisson test, we first counted the number of accumulated spikes across the trials within a 1 ms bin along the 500 ms period before stimulation. Using this, a histogram was constructed, with the abscissa representing the number of spikes and the ordinate representing the number of bins that had the number of spikes corresponding to the values on the abscissa. The histogram was fitted with a Poisson distribution curve. Using the Poisson curve, the threshold value of spikes per bin was determined that matched the *p* value of 0.01. Then, the number of spikes of each bin during poststimulation period was examined to see whether it exceeded the significance level. This Poisson method was quite effective for positive deviation (excitatory modulation), but it was not effective when the neuron had a low background firing rate, like DA neurons. Wilcoxon signed rank test was used to detect the orthodromic stimulation-induced suppression of DA neuron activity. We averaged 10 trials of spiking activity to increase the statistical reliability of the low-frequency DA firing. The modulation of orthodromic stimulation of the RMTg was performed in the same way as the one described above, except for the averaging of 10 trials.

Results

We examined the properties of LHB, DA, and RMTg neurons in two rhesus monkeys (B and C; LHB, *n* = 31; DA, *n* = 30; RMTg, *n* = 82). To locate the RMTg, we used magnetic resonance images as a guide (Fig. 1A) and advanced the recording microelectrode to the area between the VTA and the median raphe (MR), which corresponds to the RMTg in the rat (Jhou et al., 2009a). The electrode penetration was tilted laterally by 35° (Fig. 1B) and thus allowed us to explore the areas near the midline including the VTA and the median raphe (Fig. 1C). Firing properties of neurons along the electrode penetrations gave us important clues to localizing the recording sites. Particularly important were eye movement-related activities in the oculomotor (third) nucleus, the trochlear (fourth) nucleus, and the nucleus reticularis tegmenti pontis. Also useful were prominent fiber tracts including the medial longitudinal

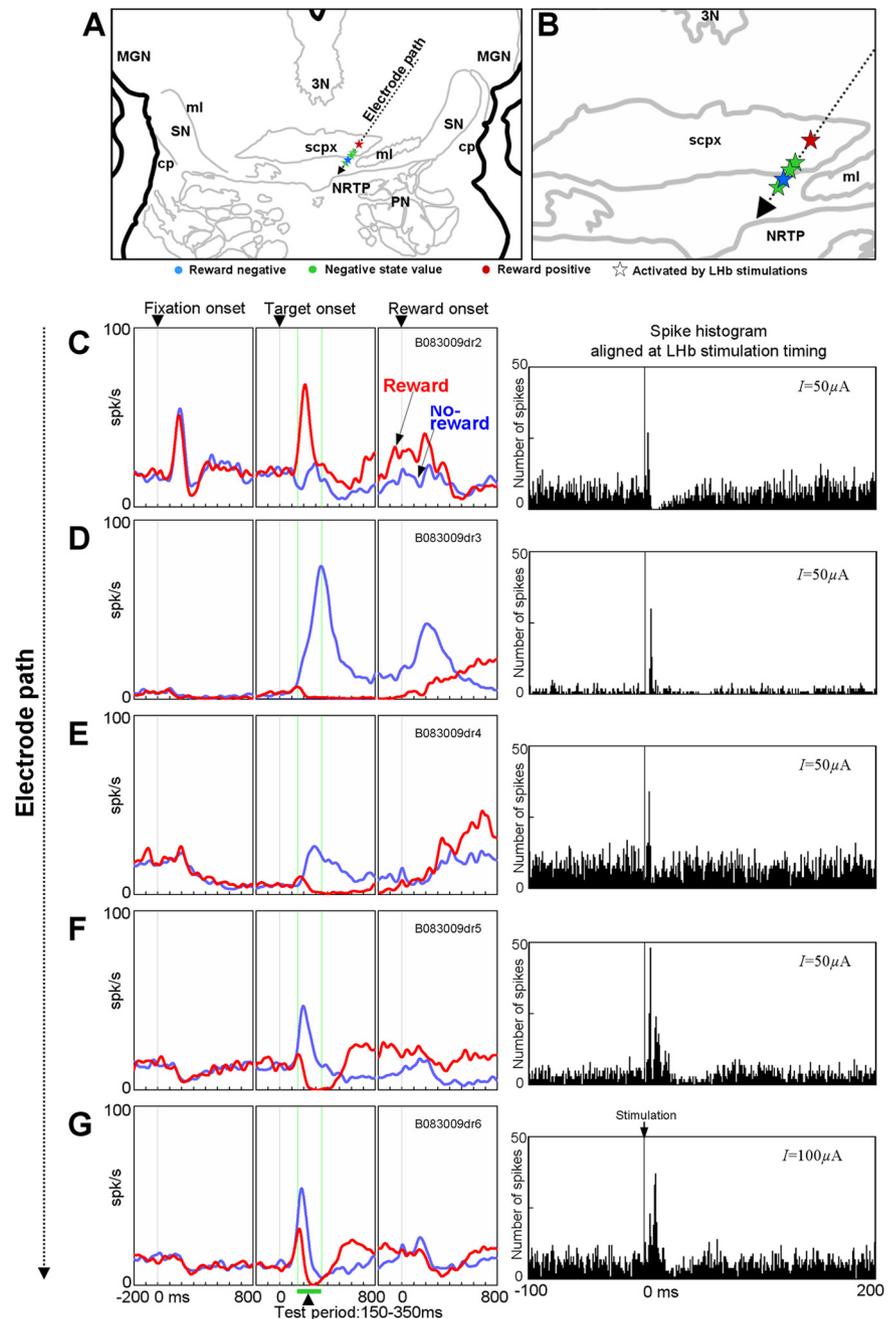


Figure 4. Sample electrode penetration aimed at the RMTg. Spike activity was recorded from five neurons along the penetration. **A**, Reconstructed positions or neurons. **B**, Magnified part around the recording positions in **A**. For each neuron located from top to bottom, the activity in the 1DR task is shown on the left in **C–G**, separately for reward trials (red) and no-reward trials (blue). The neuronal activity is aligned at the time of the fixation point onset (left), target onset (center), and reward onset (right; gray vertical lines). The green lines indicate the time window used to quantify the neuronal response to the target onset (150–300 ms after target onset). The effect of the LHB electrical stimulation is shown on the right side for each neuron as a perstimulation time histogram (bin width, 1 ms). The stimulation current was 50 μ A for neurons in **C–F**, and 100 μ A for neurons in **G**. 3N, Third nucleus; cp, cerebral peduncle; PN, pontine nuclei; NRTg, nucleus reticularis tegmenti pontis; ml, medial lemniscus; MGN, medial geniculate nucleus; scp, superior cerebellar peduncle decussation; SN, substantia nigra.

fasciculus and the superior cerebellar peduncle decussation (scp).

To examine reward-related properties of recorded neurons, we had the monkey continuously perform the 1DR task (Fig. 2A) while advancing the electrode. Both monkeys showed significantly shorter saccade latencies in rewarded trials than in unrewarded trials (Fig. 2B), indicating that they had learned the position-reward contingencies (Hikosaka et al., 2006).

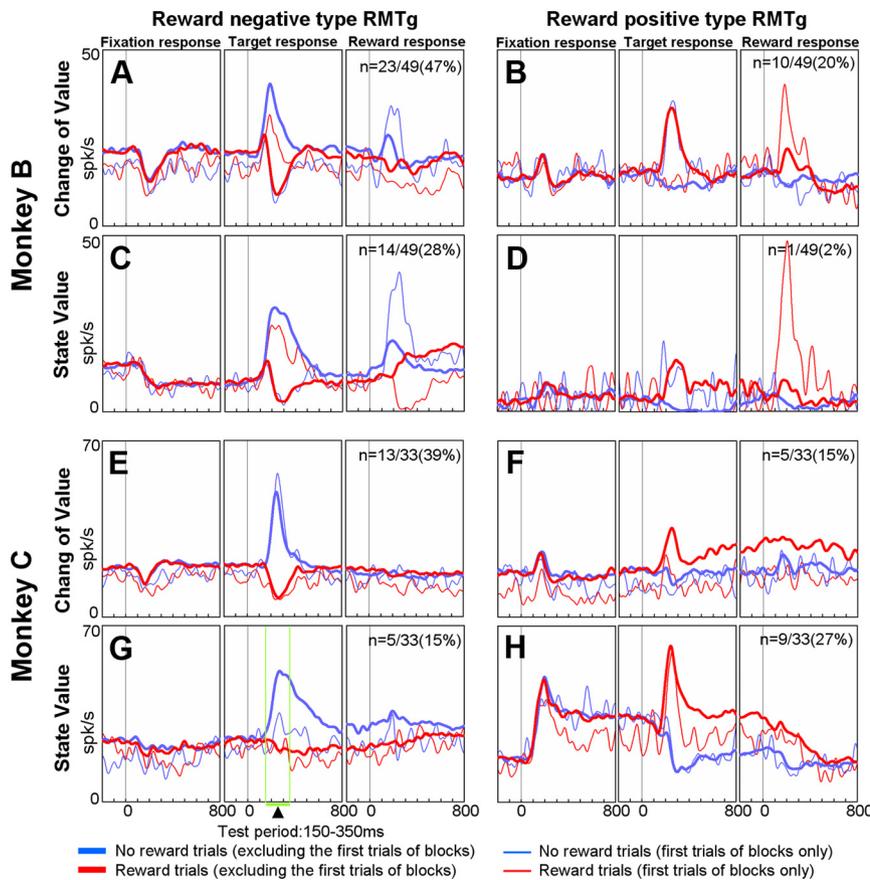


Figure 5. A–H, Classification of RMTg neuron types. For each monkey (B and C), the averaged neuronal activity in the 1DR task is shown separately for reward-negative neurons (left) and reward-positive neurons (right). Each group of neurons is further classified into the change-of-value type (top) or the state value type (bottom). The neuronal activity is aligned at fixation point onset, target onset, and reward onset. The thin SDFs indicate the average activity in the first trials after the reversal of the position-reward contingency, while the thick SDFs indicate the average activity except for the first trials. The average activity is shown separately for the reward trials (red) and no-reward trials (blue).

We found many reward-related neurons in the RMTg. An example is shown in Figure 3B. The neuron was excited after the onset of the target predicting nonreward (blue) and was inhibited after the onset of the target that predicted upcoming reward (red). This was true regardless of the location of the target. This negative-reward-modulation feature was similar to that of LHB neurons (Fig. 3A), but was opposite to that of DA neurons, which showed positive reward modulation (Fig. 3C).

Reward-related properties of primate RMTg neurons

Figure 4 shows five reward-related neurons recorded along a single track of penetration through the RMTg. All of these neurons responded to the target differentially depending on its reward value indicating reward (shown in red) and no reward (blue). The first neuron encountered in this penetration was of the “reward-positive” type (excited by the reward-indicating target) (Fig. 4C). This neuron was embedded within the lateral portion of the decussation of the scpX, dorsolaterally to the main cluster of RMTg neurons recorded in this study. The subsequent four neurons were of the “reward-negative” type (excited by no-reward-indicating target) (Fig. 4D–G) and were located within the main cluster of recorded RMTg neurons. Although broadly classified as reward negative, these neurons showed some differences in activity patterns. The two ventralmost neurons on this track (Fig. 4F,G) exhibited phasic changes mostly restricted to a 150–350

ms window after stimuli, with inhibitions after reward-predictive cues and excitations after targets predicting nonreward. These responses are similar to the negative reward prediction errors seen in LHB neurons, in which transient changes in firing represent the values of instantaneous changes in predicted reward. However, the two more dorsal neurons on this track (Fig. 4D,E) showed sustained decreases in activity after both the trial start event (fixation point onset) and the saccade-target cue indicating reward or no reward. These responses extended into a 350–700 ms window after stimuli onset, and even longer in some instances. These neurons’ persistent firing changes are consistent with representations of state value (Belova et al., 2008; Bromberg-Martin et al., 2010a), approximately corresponding to the levels (or states) of the expected reward values, rather than their instantaneous change. Notably, one neuron (Fig. 4G) exhibited both types of responses, with a small but significant ($p < 0.01$) tonic representation of state value on top of the phasic response.

To summarize, the reward-related neurons in the RMTg were classified into two groups (the reward-positive type and the reward-negative type), and each group was classified into two subgroups (change-of-value type and state-value type). The averaged activity for each of the four groups is shown in Figure 5, separately for the two monkeys. Of the 82 RMTg neurons that we recorded, 55 neurons (67%) were of reward-negative type, and 25 neurons (30%) were of reward-positive type. Among the reward-negative type, 36 neurons (65%) were of change-of-value type, and 19 (35%) were of state-value type. Among the reward-positive type, 15 neurons (60%) were of change-of-value type, and 10 (40%) were of state-value type. Two neurons were left unclassified because their responses occurred after the test time window (150–350 ms after target).

In addition to the responses to the target, the RMTg neurons also responded to the onset of the fixation point and the reward outcome. The onset of the fixation point induced an inhibition in the reward-negative neurons and an excitation in the reward-positive neurons (Fig. 5). This type of response is consistent with prediction error encoding, as the fixation point initiates a new trial and signifies an increased likelihood of obtaining an impending reward. Similar responses to the fixation point were present in LHB neurons and DA neurons (Fig. 6), as reported previously (Bromberg-Martin et al., 2010b). Note that the response to the onset of the fixation point was transient in some of the change-of-value RMTg neurons and sustained in the state-value neurons.

The reward outcome (presence or absence of reward) also evoked responses in the RMTg neurons, but clear responses were present only on the first trial after the position-reward contingency was reversed (when the reward outcome was unexpected). When monkey B had been expecting a reward but there was no reward, the reward-negative neurons increased their activity, and

the positive reward-neurons decreased their activity (Fig. 5A–D, thin blue lines). When the monkey had been anticipating no reward but there was a reward, the negative neurons decreased their activity and the positive neurons increased their activity (Fig. 5A–D, thin red lines). These results indicate that the RMTg neurons encode reward-prediction errors, similar to Lhb neurons and DA neurons (Fig. 6). There was no clear surprise outcome response in the RMTg neurons in monkey C. This monkey was trained on a variation of the 1DR task in which the color of the target indicated the presence or absence of reward, so that the outcome was always fully predicted. The results in monkey C are consistent with the hypothesis that the responses of the RMTg neurons encode reward-prediction errors.

Trial-to-trial changes in neuronal and behavioral responses

The reward-related differential responses of the RMTg neurons suggest that these neurons might contribute to the generation of the differential saccade latencies (Fig. 7C,F); namely, because the rewarded side was reversed after a block of 24 trials (e.g., right side rewarded to left side rewarded), the difference in saccade latency was reversed accordingly (e.g., right earlier than left to left earlier than right).

Figure 7, A, B, D, and E, shows the time courses of the changes in the activity of Lhb, RMTg, and DA neurons, as well as the changes in saccade latency (C, F) within a block of trials. In monkey B, all groups of neurons (Lhb, reward-positive RMTg, reward-negative RMTg, and DA neurons) showed similar changes in activity. All of them reversed their responses to the target onset (post-target) between the first trial and the subsequent trials. On the first trial of the block, the response patterns are similar to those of the preceding block because the monkey did not yet know that the rewarded target position has changed. On the second and subsequent trials, the neuronal responses reversed and approached the differential pattern based on the updated position-reward contingency. The red line shows the transition of the neuronal response from the unrewarded to rewarded condition and the blue line from the rewarded to unrewarded condition. The saccade latency showed a similar reversal: it decreased quickly after the transition from the unrewarded condition to the rewarded condition (Fig. 7C, red line), but increased slowly after the transition from the rewarded to unrewarded condition (Fig. 7C, blue line) (Hikosaka et al., 2006). In contrast, monkey C showed little change in saccade latency (Fig. 7F) because the change of the rewarding position was predictable by the colored target even on the first trial of the block. In other words, the saccade latency changed abruptly on the trial when the color of the target changed, indicating the reversal of the rewarded position. Accordingly, the neuronal responses changed abruptly on the first trial and remained basically unchanged on the subsequent trials.

The neuronal responses to the reward outcome (postreward) (Fig. 7B), as measured by the difference between the activity in the

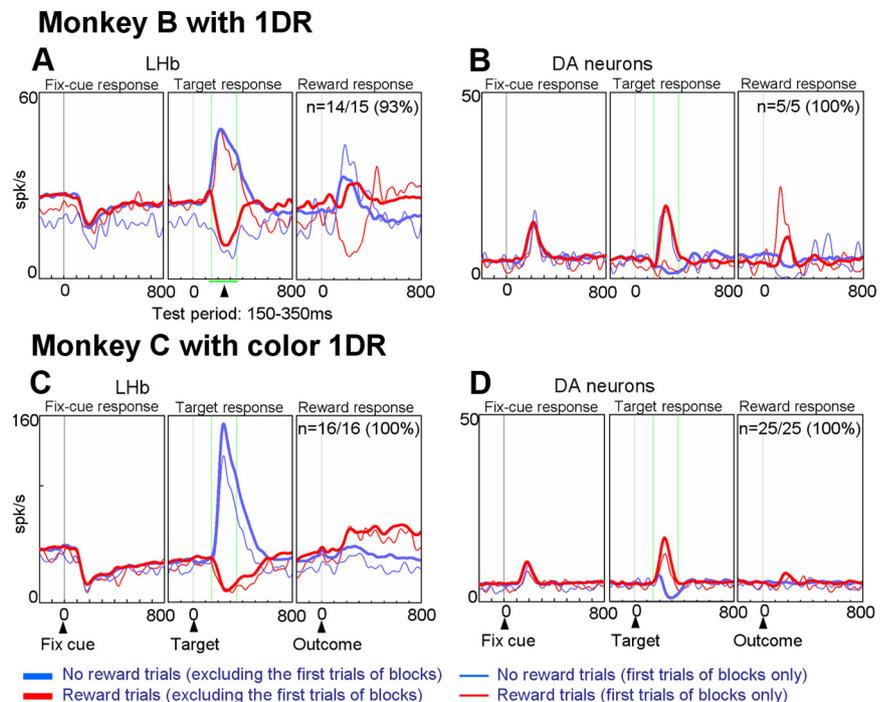


Figure 6. Population activity of Lhb and DA. **A–D**, Population responses of Lhb neurons (monkey B, $n = 14$, **A**; monkey C, $n = 16$, **C**) and DA neurons (monkey B, $n = 5$, **B**; monkey C, $n = 25$, **D**). The SDFs are shown for each reward contingency (red, rewarded trials; blue, unrewarded trials). Thick SDFs indicate activity excluding the first trial in each block. Thin SDFs indicate the average activity of the first trial of the block. The two vertical green lines show the time window used to test these responses (reward positive or reward negative). The analysis in **C** contains some multiunit recordings.

rewarded condition (red line) and the activity in the unrewarded condition (blue line), were exclusive to or largest on the first trial, again consistent with encoding of reward-prediction errors. However, the difference in activity between the rewarded and unrewarded conditions remained on subsequent trials, especially among RMTg neurons compared with Lhb or DA neurons. This may be related to the fact that some of the RMTg neurons showed activity related to the state value. In summary, in both monkey B and monkey C, neuronal response changes paralleled saccade latency changes.

Electrophysiological parameters of different groups of neurons

Since the primate RMTg is unknown, we characterized the neural parameters of the RMTg, including spike shape, irregularity of firing, and latencies of orthodromic and antidromic activation (Fig. 8). The mean spike width (the length between the two troughs before and after the peak of the spike) of the RMTg neurons was 0.66 ms (SD, ± 0.17 ms) with a mean baseline firing rate of 17.8 (± 12.1) Hz ($n = 82$). Their average irregularity in spike timing (Davies et al., 2006) was 1.081 (± 0.421). Statistical analysis shows that the spike widths of the RMTg neurons were marginally wider than those of Lhb neurons (0.55 ± 0.14 ms, $0.01 < p < 0.05$). The spike widths of DA neurons were wider than those of Lhb and RMTg neurons ($p < 0.01$), and their baseline firing frequency was low (~ 5 Hz) (Fig. 8), conforming to the known features of DA neurons (Matsumoto and Hikosaka, 2009).

Determination of RMTg connectivity using orthodromic and antidromic stimulations

We next tested the hypothetical connectivity, Lhb \rightarrow RMTg \rightarrow SNC/VTA, using electrophysiological and anatomical methods. The first electrophysiological method was orthodromic stimula-

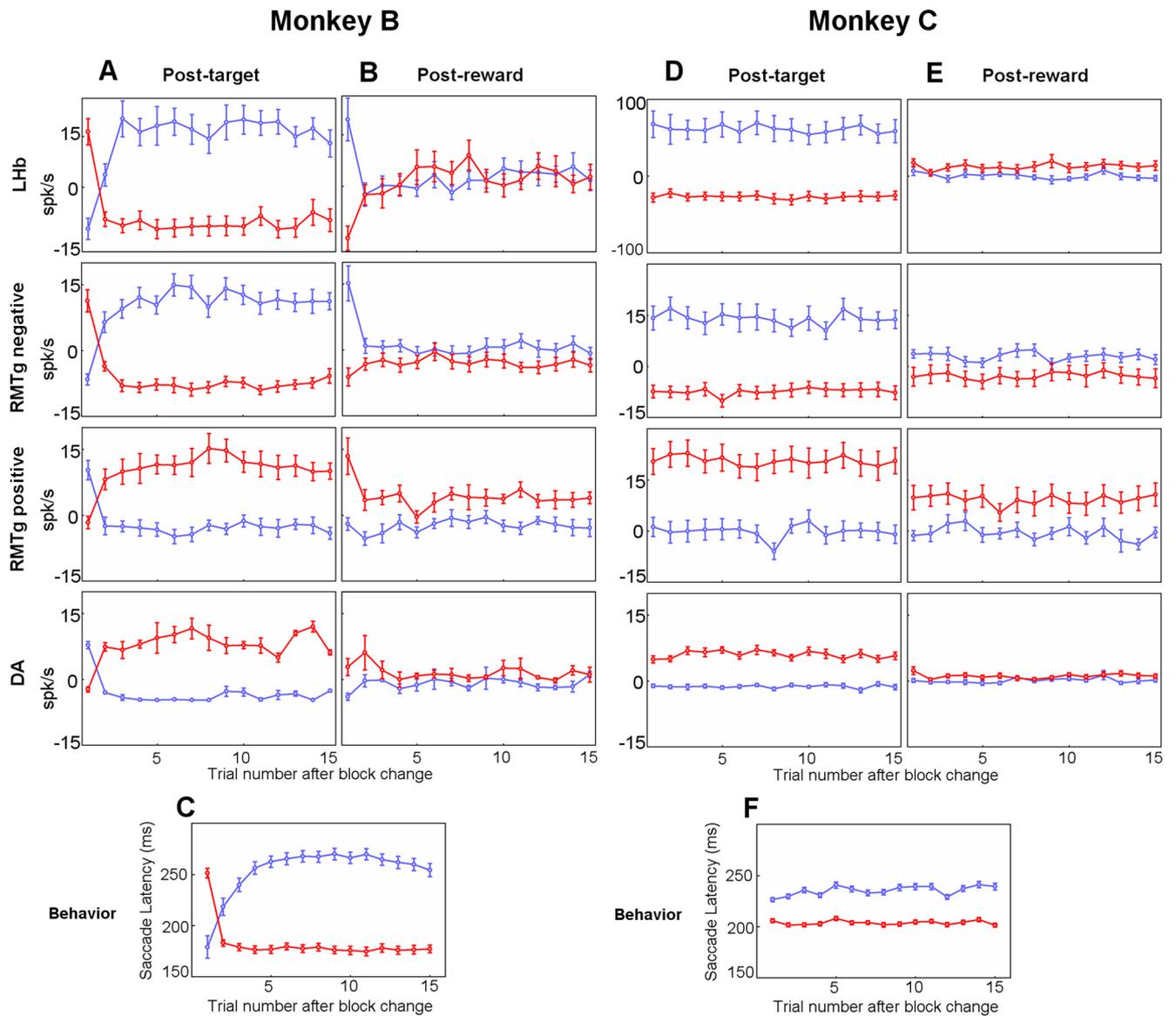


Figure 7. Within-block changes of neural responses and behavioral (saccade) latencies. **A–F**, Changes in baseline subtracted averaged post-target responses (**A, D**), averaged reward on–off responses (**B, E**), and averaged saccade latency (**C, F**) after the reversal of position–reward contingency are shown. The data from two monkeys are shown separately (left half from monkey B, right half from monkey C). Red and blue lines indicate the data in rewarded and unrewarded trials, respectively. The data from ipsilateral and contralateral saccades are combined. Error bars indicate SEM. Lhb data from monkey C contain some multiunit recordings.

tion. According to the hypothesis, electrical stimulation within the Lhb should influence RMTg neuron activity via the Lhb–RMTg synapses, and stimulation of the RMTg should affect the activity of SNc/VTA neurons.

We indeed found such effects; a representative reward-negative neuron in the RMTg in Figure 9B was excited ($p < 0.01$) (Fig. 9E, F) by electrical stimulation in the Lhb where the neuron shown in Figure 9A was recorded. Sometimes the poststimulation excitation was followed by a long-latency rebound inhibition as shown in Figure 9, E and F. Similar excitations (but no inhibitions) were evoked in a majority of reward-negative neurons (18 of 30; 60%) (Table 1). The excitatory responses are consistent with the hypothesis that Lhb neurons transmit negative reward signals to RMTg neurons by excitatory synapses. The right column of Figure 4C–G shows examples of orthodromic excitation of RMTg neurons by Lhb stimulation along a recording track.

The effect of the Lhb stimulation on reward-positive neurons was less consistent. A representative reward-positive neuron

shown in Figure 9C was inhibited by the Lhb stimulation, albeit preceded by a brief excitation (Fig. 9G; same as Fig. 4C, right column). For the neurons showing biphasic responses, the first component was taken as the primary response. Among 11 reward-positive neurons, 4 were excited and 3 were inhibited.

None of the RMTg neurons tested ($n = 41$) were activated antidromically by Lhb stimulation. This is consistent with the anatomical data showing that the connection is mostly unidirectional, from the Lhb to the RMTg (Herkenham and Nauta, 1977; Jhou et al., 2009a).

Stimulation of the RMTg suppressed activity in 16 of 17 DA neurons examined (94%; $p < 0.01$, Wilcoxon signed rank test) (Fig. 9D, H). The remaining neuron was not responsive (Table 1). These inhibitory responses are consistent with the hypothesis that the RMTg suppresses DA neurons by inhibitory synapses. None of the DA neurons were activated antidromically by the RMTg stimulation.

The effect of orthodromic stimulation could be due to the activation of axons that pass through the stimulation site, not

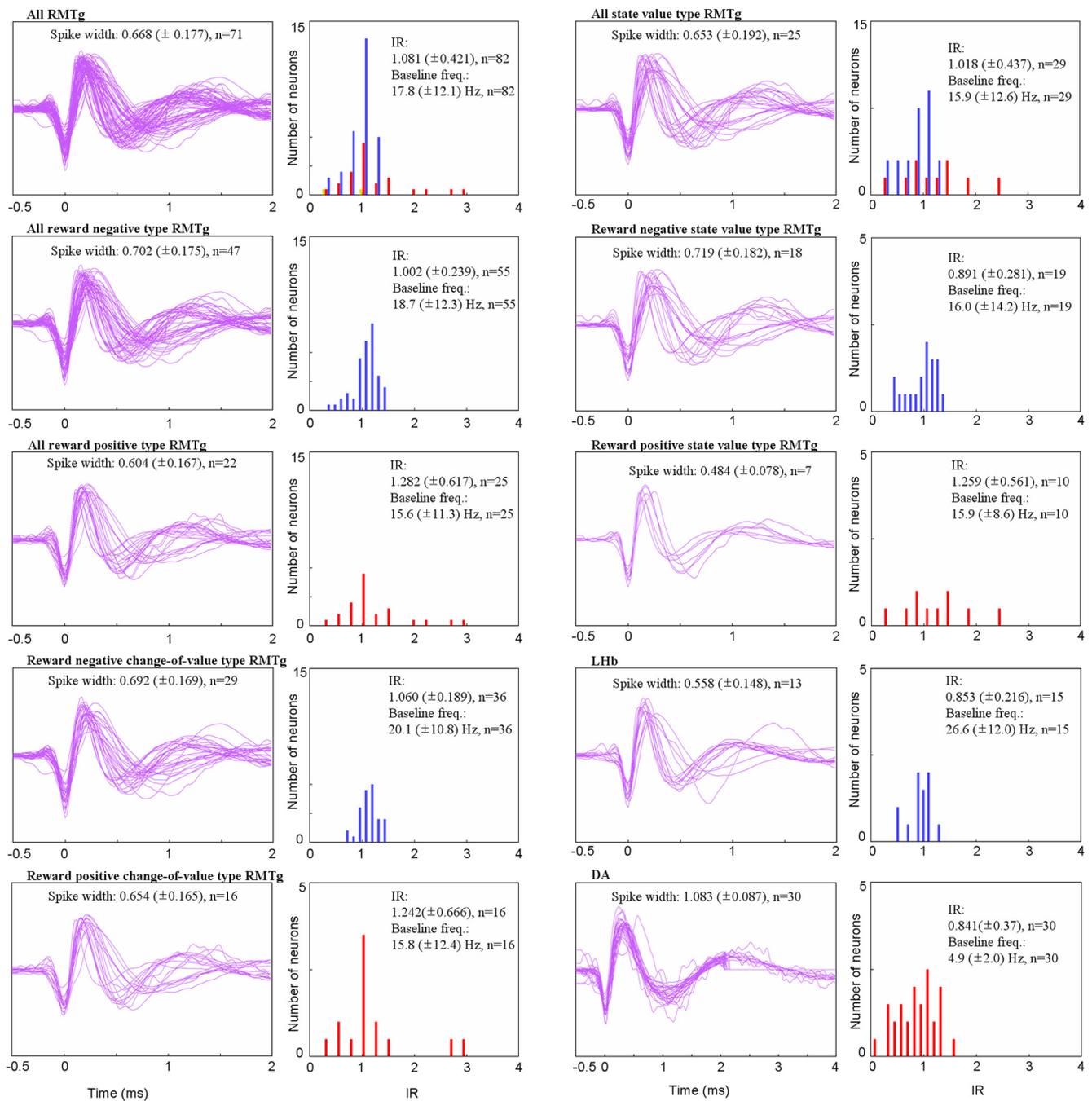


Figure 8. Average spike wavelength, irregularity index, and mean baseline firing rate of different groups of neurons. Whereas most of RMTg neuron groups showed similar average spike width to that of LHb neurons, change-of-value reward-negative neurons had a slightly longer average spike width than LHb neurons ($p < 0.05$, ANOVA). The spike widths of DA neurons are significantly longer than those of LHb neurons ($p < 0.01$, ANOVA). Note that the scale of the abscissa for DA neurons' spike width is compressed. The absence of signals after 2 ms in some DA neuron spike shapes is due to changes in our recording settings. The width of the spike is defined as the time between the two negative peaks. The average irregularity in spike timing (IR) was smallest in DA neurons followed by LHb neurons. In the histograms for "All RMTg" and "All state value type RMTg," different subclasses of the group are represented in different colors (positive, red; negative, blue; null, yellow).

activation of neuron somas located at the stimulation site. To exclude this possibility, we used antidromic stimulation. If LHb neurons, not passing axons, project to the RMTg, electrical stimulation of the RMTg should activate the LHb neurons antidromically. We conducted such an antidromic stimulation experiment in two recording sessions and found two antidromically activated neurons in the LHb that were activated by the RMTg stimulation. One example neuron is shown in Figure 10A. The antidromic activation latency was 4 ms, and the threshold current for the

activation was very low ($10\mu\text{A}$) (Fig. 10D). Table 2 summarizes the electric stimulation parameters between structures.

Applying the same method to the RMTg-DA connectivity, we found four RMTg neurons antidromically activated by the stimulation of SNC. We first recorded from a typical DA neuron in the SNC, which showed reward-positive responses in the 1DR task (Fig. 10C). Having finished the recording of this DA neuron, we stimulated this location and found an antidromically activated neuron in the RMTg (Fig. 10E). This RMTg neuron turned out to

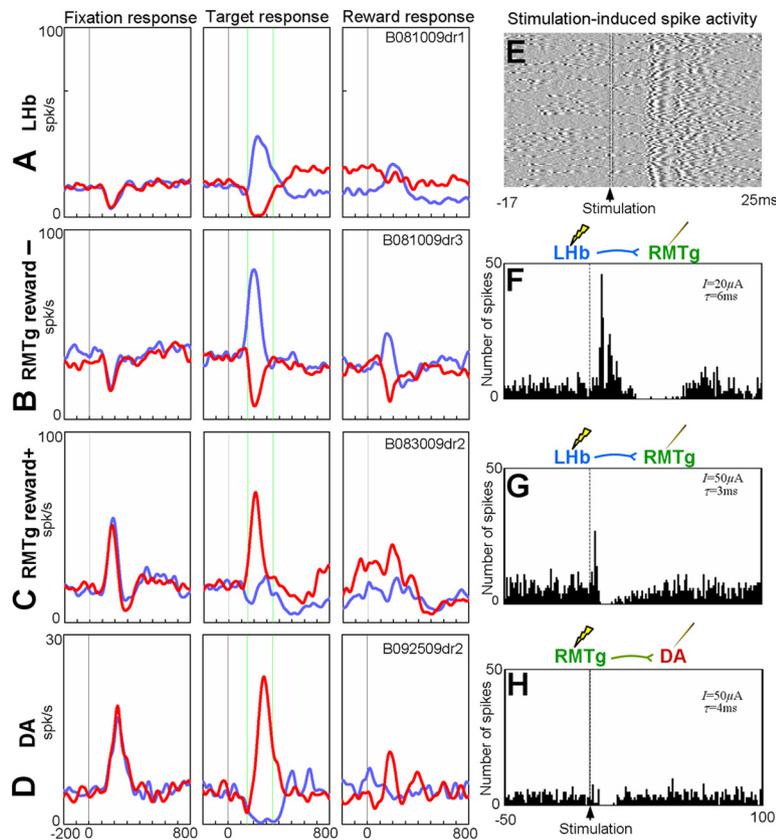


Figure 9. Orthodromic responses along the Lhb–RMTg–DA circuit. **A–D**, Averaged activity of a single neuron in each area during the 1DR task: an Lhb neuron (**A**), a reward-negative RMTg neuron (**B**), a reward-positive RMTg neuron (**C**), and a DA neuron (**D**). **E–H**, Orthodromic responses. **E, F**, Response of the reward-negative RMTg neuron (**B**) to the electrical stimulation in the Lhb. **G**, Response of the reward-positive RMTg neuron (**C**) to the stimulation in the Lhb. **H**, Response of the DA neuron (**D**) to the stimulation in the RMTg. In **F–H**, the orthodromic responses are shown as peristimulus time histograms (bin width, 1 ms). In **E**, the orthodromic response to the Lhb stimulation is shown as the actual voltage changes (negative, black; positive, white) associated with the extracellular action potentials of the reward-negative RMTg neuron (**B**). The yellow lightning bolt symbols in **F–H** indicate the stimulation site; the sharp needle shape indicates the recording site. *I* indicates the threshold current, and τ indicates the latency for the orthodromic response.

Table 1. Summary of orthodromic stimulation effects

Stim Lhb → Record RMTg	RMTg Neuron types		
	Negative	Positive	All
Excited	18	4	22 (52%)
Inhibited	0	3	3 (7%)
No modulation	12	4	17 (40%) ^a
All	30 (71%)	11 (26%)	42
Stim RMTg → Record DA	DA Neuron types		
	Negative	Positive	All
Excited	0	0	0 (0%)
Inhibited	0	16	16 (94%)
No modulation	0	1	1 (6%)
All	0 (0%)	17 (100%)	17

^aOne RMTg neuron could not be classified as negative or positive.

be of reward-negative type (Fig. 10B). This result suggested that the RMTg neuron exerted an inhibitory effect on the DA neuron. Indeed, when we stimulated this location of the RMTg neuron, the DA neuron was inhibited (Fig. 10F). As a further test of this RMTg→SNc projection, we used this SNc site to inject a retrograde tracer for an anatomical study (see Retrograde tracing study).

The particular result described above is consistent with the hypothesis that reward-negative information is sent from the RMTg to DA neurons in the SNc. Interestingly, however, two of the four antidromically activated RMTg neurons were of reward-positive type (Fig. 11, brown dots with black contour). In those cases, we could not find well-isolated DA neurons at the antidromic stimulation site and therefore could not test the effect of orthodromic stimulation from these sites.

Neuronal reward response latency

The orthodromic and antidromic stimulation experiments described above suggest that reward-related information is sent from the Lhb to the RMTg via largely excitatory synapses, and then from the RMTg to DA neurons in the SNc/VTA via inhibitory synapses. If so, the latency of the reward effect after target onset is predicted to be shortest among Lhb neurons, followed by RMTg neurons, and finally DA neurons. To test this prediction, we examined the latency at which each of these groups of neurons started differentiating their activity depending on the expected reward outcome. Figure 12 shows the percentage of the neurons that differentiated the reward/no-reward contingency after target onset (for details, see Materials and Methods). The differentiation latency was determined as the time after target onset when the ratio of the neurons showing reward/no-reward differentiation exceeded the criterion level of significance (Fig. 12, horizontal gray line; $p = 0.01$). The determined differentiation

latencies were 141 ms for reward-positive RMTg neurons and 136 ms for reward-negative RMTg neurons, which were both shorter than the latencies of 147 ms for Lhb neurons and 161 ms for DA neurons. However, a close comparison of the latencies (Fig. 12E) indicates that, although a small number of RMTg neurons started the differentiation earlier than Lhb neurons, the differentiation of other RMTg neurons tended to lag behind Lhb neurons. These results suggest that the later part of the reward-related activity of RMTg neurons could be derived from the Lhb.

Anatomical study of primate RMTg

We determined the locations of these reward-related neurons anatomically. This anatomical study is divided into two parts: (1) histological reconstruction of the locations of the reward-related neurons and (2) a retrograde tracer study of the RMTg–SNc connection.

We first made electrolytic lesions in the RMTg on the left and right sides of the brain stem where the reward-related neurons were recorded (Fig. 13B, arrows). These lesions (one pair on each side of the brain stem) were made at the top and bottom parts of the presumed RMTg along the recording electrode track by passing a negative current of 13 μ A for 30 s. The right-side pair of lesions is visible in the section shown in Figure 13B; the left-side pair of lesions is visible in slightly posterior sections (Fig. 13E, F), but has also been

projected onto the section in Figure 13B (two arrows on the left side of the brain). Recording sites at two anterior–posterior levels separated by 1 mm have been combined.

Based on marking lesions described above, our reconstruction of reward-related neurons in monkey B is shown in Figure 13C. Reward-related neurons were clustered in the area lateral and dorsal to the interpeduncular nucleus (IP), largely below the scpx, analogous to the RMTg location in rats (Jhou et al., 2009a). In monkey C, we explored a larger area, particularly more posteriorly. At this level, reward-related neurons were clustered lateral to the MR similarly to rats (Fig. 11).

It was notable that a high density of reward-related neurons was found below the scpx, bilaterally. Within the central regions of the RMTg (Fig. 13C, 1 mm diameter circles on both sides of the RMTg bound by a pair of arrows), more reward-related neurons were reward-negative types (30 of 32, or 94%; this count includes only cells visible in Fig. 13C where the RMTg center can be defined clearly) compared to the neurons outside of these central zones (10 of 16, or 63%; $p < 0.006$; Pearson's χ^2 test). A similar pattern of diffuse boundaries of the RMTg has been observed in rats (Jhou et al., 2009b).

Retrograde tracing study

The second part of the anatomical investigation was a retrograde tracer study of the RMTg–SNc connection. At the conclusion of the single-unit recording study, we injected CTB (0.9%, 0.2 μ l) into the right SNc (Fig. 14). This injection site was estimated to be the same location where the DA neuron shown in Figure 10C was recorded. Selection of this SNc site is meaningful because we observed bidirectional stimulation effects (orthodromic and antidromic) between this SNc site and the RMTg, as shown in Figure 10, E and F.

Retrogradely labeled neurons were found throughout the midbrain (Fig. 13D), but also with a distinct cluster occurring around the electrolytic lesions (arrows), where many reward-related neurons were recorded (Fig. 13C). This cluster resided largely ventral to the scpx and was most prominent ipsilaterally, with a similar but weaker pattern present contralaterally (Fig. 13D). A second cluster of retrogradely labeled neurons was seen just dorsal to the scpx near the midline (arrowhead). Although we did not survey extensively, we also found a few reward-related neurons around that region, as shown in Figure 13C. Notably, the RMTg in rats also has a vertical elongation in some sections (Jhou et al., 2009b; Kaufling et al., 2009), raising the possibility that this elongation has progressed further in primates.

The good match between the centers of the RMTg neuron areas deduced from our electrophysiological method (demarcated by the marking lesions) and the anatomic locations of the SNc-projecting neurons supports our electrophysiological findings.

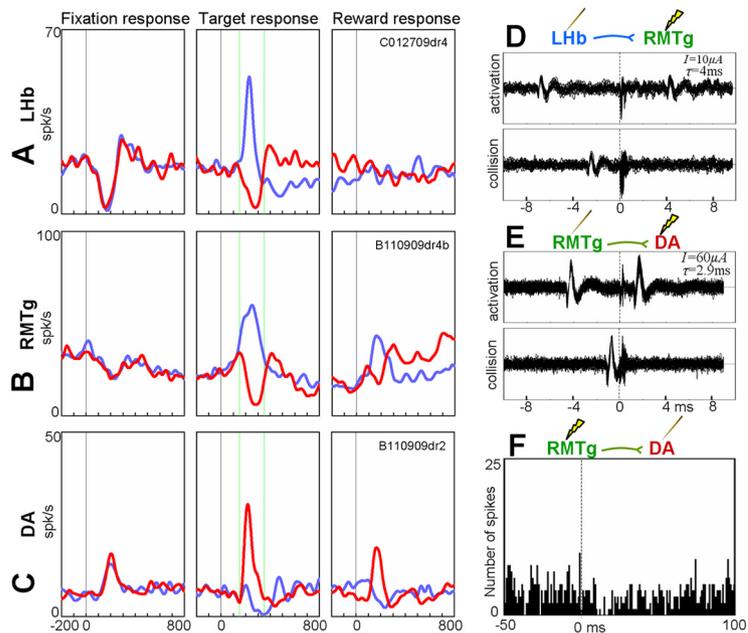


Figure 10. Antidromic responses along the Lhb–RMTg–DA circuit. **A–C**, Averaged activity during the 1DR task for an Lhb neuron (**A**), a reward-negative RMTg neuron (**B**), and a DA neuron (**C**). The same format as Figure 5 is used. **D**, Antidromic responses of the Lhb neuron to the electrical stimulation in the RMTg. The stimulation was delivered with a fixed time delay after the spontaneous spike of the Lhb neuron. Antidromic spikes occurred when the delay was long enough (top), but was blocked due to collision when the delay was shorter (bottom). **E**, Antidromic responses of the RMTg neuron to the stimulation in the SNc where the DA neuron (**C**) was recorded. Conversely, stimulation at the recording site of the RMTg neuron induced an inhibition in the DA neuron (**F**). Note that the CTB injection site shown in Figure 14 was aimed at the recording site of this DA neuron.

Table 2. Summary of electric stimulation parameters

	Orthodromic	Antidromic
Lhb–RMTg threshold current	75 ± 51 μ A	10 μ A, 50 μ A; $n = 2$
Lhb–RMTg latency	4.8 ± 1.8 ms	2.6 ms, 4 ms; $n = 2$
RMTg–SNc threshold currents	156 ± 60 μ A	62 ± 55 μ A; $n = 4$
RMTg–SNc latency	4.4 ± 1.3 ms	2.1 ± 0.9 ms; $n = 4$

Discussion

Localization of RMTg in the monkey

This study localized and characterized tegmental neurons in macaque monkeys that transmit reward-related information from the Lhb to DA neurons in the SNc. We found that such reward-related neurons were localized in the paramedian tegmental area, caudal to the VTA (Fig. 13), extending caudally toward the pedunculopontine tegmental nucleus (PPTg) along the lower border of the scpx (Fig. 11). This location is remarkably similar to the tegmental neurons in the rat that were shown anatomically to receive inputs from the Lhb (Herkenham and Nauta, 1979; Jhou et al., 2009a; Kaufling et al., 2009; Kim, 2009) and send outputs to the SNc/VTA (Jhou et al., 2009a; Kaufling et al., 2009). The similarity is particularly evident at its rostral border (Fig. 13), where it is located just lateral to the IP and below the scpx.

The RMTg as a mediator of Lhb-induced inhibition of DA neurons

We found that a majority of neurons in the RMTg were of reward-negative type and were insensitive to the position of the target. This activity pattern was similar to that of Lhb neurons. Many of the reward-negative RMTg neurons were excited, and none were inhibited, by electrical stimulation of the Lhb. Some Lhb neurons were activated antidromically from the RMTg, suggesting that the excitation was mediated by direct connections

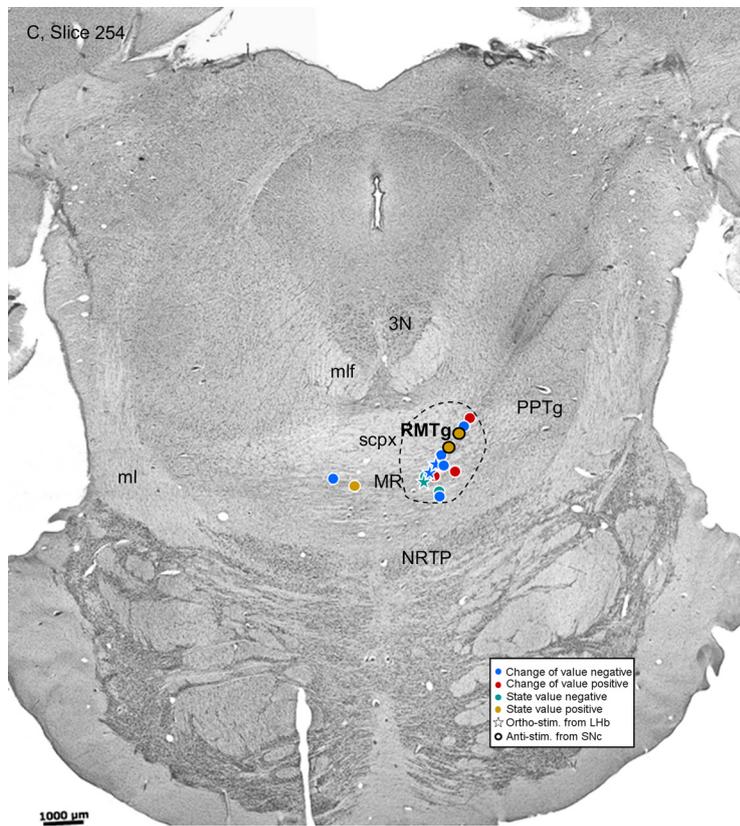


Figure 11. Estimated locations of the recorded neurons in the RMTg in monkey C. NRTg, Nucleus reticularis tegmenti pontis; ml, medial lemniscus; 3N, third nucleus; mlf, medial longitudinal fasciculus; PPTg, pedunculopontine tegmental nucleus; scpX, superior cerebellar peduncle decussation; MR, median raphe.

from Lhb neurons to RMTg neurons. Electrical stimulation of the RMTg in turn induced an inhibition in putative DA neurons in the SNc, and some RMTg neurons were activated antidromically from the SNc. These results are consistent with recent observations in the rat indicating that Lhb neurons are excitatory (Geisler and Trimble, 2008; Omelchenko et al., 2009; Brinschwitz et al., 2010) and project to the RMTg (Herkenham and Nauta, 1979; Jhou et al., 2009a; Kaufling et al., 2009; Kim, 2009), and that RMTg neurons are GABAergic and inhibitory (Kirouac et al., 2004; Perrotti et al., 2005; Olson and Nessler, 2007; Jhou et al., 2009a; Kaufling et al., 2009) and project to the SNc and VTA (Jhou et al., 2009a; Kaufling et al., 2009). Particularly, a recent study by Balcita-Pedicino et al. (2011) showed that Lhb axons in the RMTg area preferentially (>55%) contact GABAergic neurons. They also showed that the GABAergic RMTg axons in the VTA contacted dendrites immunoreactive for the DA synthetic enzyme tyrosine hydroxylase. Going one step further by demonstrating both Lhb→RMTg projections and RMTg→DA neuron projections in the primate, our data represent the first demonstration of functional connectivity in this pathway.

It is thus likely that reward-negative information is transmitted from the Lhb to the RMTg via excitatory connections, and that reward-negative information in the RMTg is translated into reward-positive information in DA neurons in the SNc/VTA via inhibitory connections. Most RMTg neurons encoded reward-prediction errors, i.e., the difference (or change) between the expected reward value and the actual reward value, similarly to Lhb neurons and DA neurons in the SNc/VTA. Thus, the Lhb–RMTg–SNc/VTA circuit is likely a prominent source of the reward-prediction error signals in DA neurons. However, we also found several unforeseen results that allowed us to extend our original hypotheses, as described in the following three sections.

RMTg may receive reward-related signals from areas outside the Lhb

Our analysis indicated that the reward-related activity started earlier in some RMTg neurons than in Lhb neurons, although other RMTg neurons followed Lhb neurons (Fig. 12). This suggests that the RMTg receives reward-related information from areas other than the Lhb in addition to the Lhb. Previous anatomical studies using rats showed that the RMTg

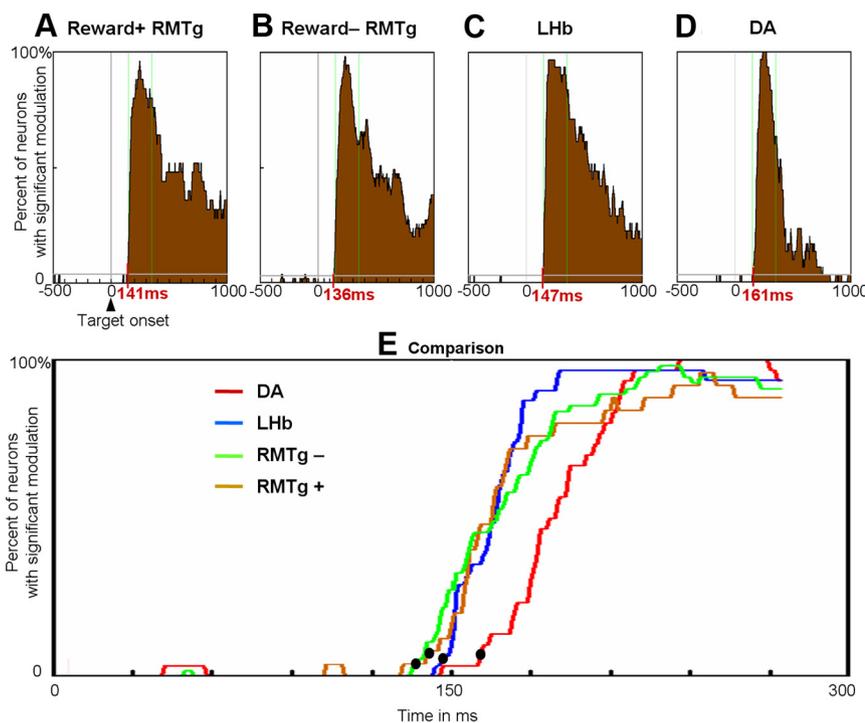


Figure 12. Neural latency of reward/no-reward discrimination. **A–D**, A time-varying proportion of neurons that showed significantly different activity between rewarded trials and unrewarded trials for the four neuron groups indicated above each figure. The horizontal gray line in each panel indicates the criterion level ($p = 0.01$). The red vertical bar and corresponding time in each panel is the time point of significant deviation from the background. Time 0 indicates target onset. For details, see Materials and Methods. **E**, Comparison of the reward-differential latency in different groups of neurons. The parts from 0 to 300 ms from each of the neural groups (**A–D**) are blown up and superimposed. The time point of significant deviation from the background for each group is indicated as a dot.

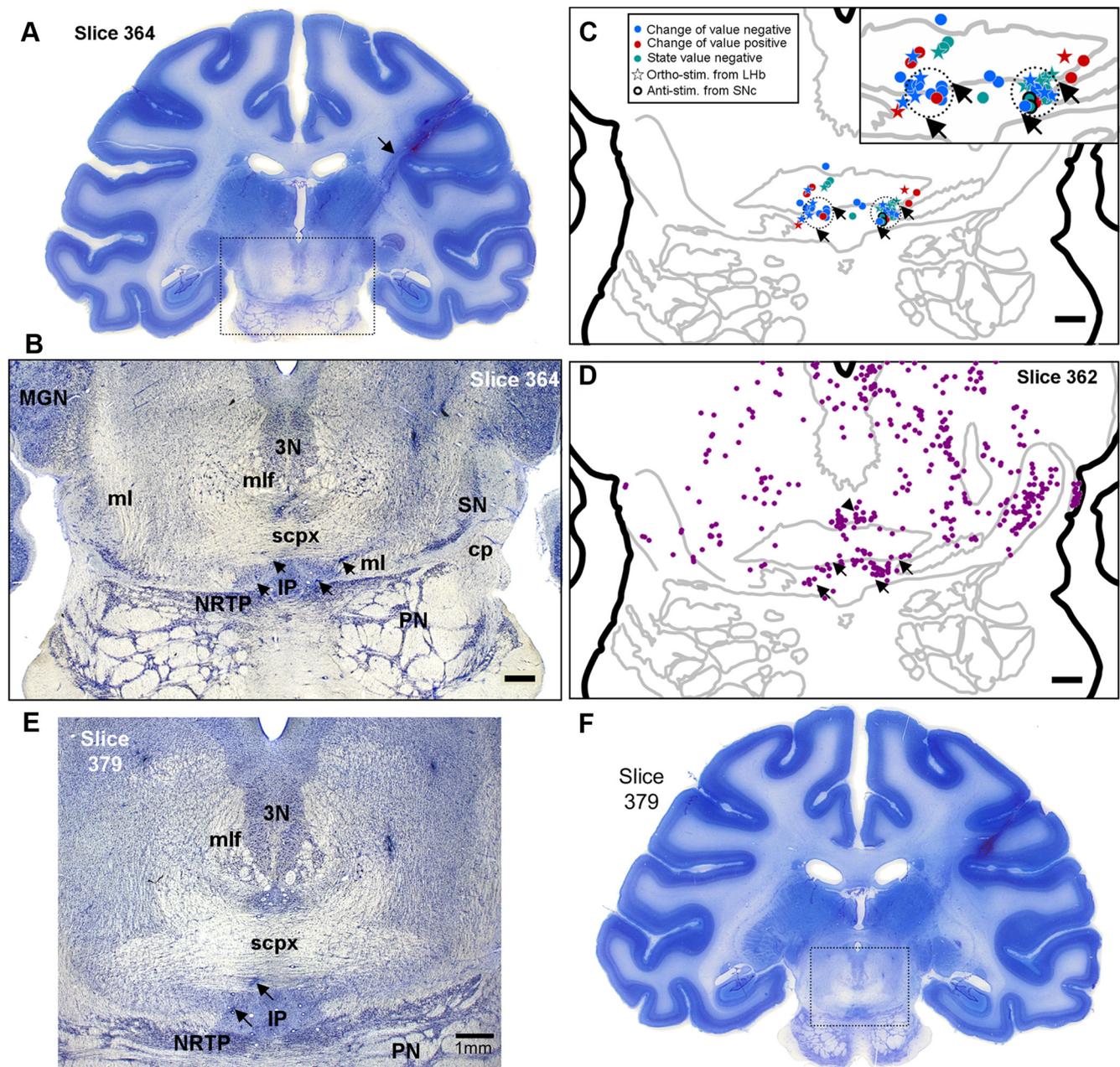


Figure 13. Identification of RMTg. **A**, The coronal brain slice showing the RMTg and other neighboring structures in the midbrain/pons. Note that some of the electrode tracts are visible on the top right part of the slice (arrow) approaching the target at about 35° angle. **B**, Magnified part of the brain section shown in **A**. This histology section shows two electrolytic marking lesions on the right side of the brain under the scpx (the two rightmost arrows). The lesions were made at the conclusion of the recording to demarcate the dorsal and ventral aspects of the RMTg along the recording tract. The leftmost arrows indicate the corresponding points of marks that are visible on a nearby slice shown in **E** and **F**. They have been projected onto this section. **C**, Estimated sites where we recorded RMTg neurons around slice 364 shown in **B** (the inset shows a magnified portion of the recording sites). We found numerous RMTg neurons responding to reward-related events under the scpx and beside the interpeduncular nucleus. Many of those neurons showed orthodromic responses to the stimulation in the LHb (stars). Some of them also showed antidromic responses to the stimulation in the SNc (2 black circles on the right side). Recording sites at two anterior–posterior levels separated by 1 mm have been combined. Marking lesion sites are indicated by arrows. **D**, Retrogradely labeled neurons after injection of CTB in the SNc. A cluster of retrogradely labeled neurons was found at the site of the RMTg. Another cluster of labeled neurons was found above the scpx close to the midline (arrowhead). See Figure 14 for the CTB injection site. **E**, In a section 0.75 mm posterior to that in **A** and **B**, two electrolytic marking lesions are visible on the left side of the brain (arrows). Scale bars: **B–D**, 1 mm. cp, Cerebral peduncle; ml, medial lemniscus; mlf, medial longitudinal fasciculus; MGN, medial geniculate nucleus; NRTP, nucleus reticularis tegmenti pontis; PN, pontine nuclei; 3N, third nucleus; IP, interpeduncular nucleus; scpx, superior cerebellar peduncle decussation; SN, substantia nigra.

receives inputs from many brain areas other than the LHb, including the medial frontal cortex, hypothalamic areas, and ventral pallidum (VP) (Jhou et al., 2009a; Kauffling et al., 2009). A recent finding from our laboratory showed (Tachibana and Hikosaka, 2009) that many neurons in the VP show reward-dependent modulations, typically in a reward-positive manner and sometimes earlier than LHb neurons. If the input from the VP

were largely inhibitory, this could account for the rapid appearance of reward-negative signals in some RMTg neurons (Fig. 15). If this is true, the RMTg is not merely a mediator of the LHb–DA inhibition, and instead is a station where multiple reward-related signals are integrated before being sent to the SNc/VTA. In other words, the LHb and the RMTg not only share common functions, but may also have different roles.

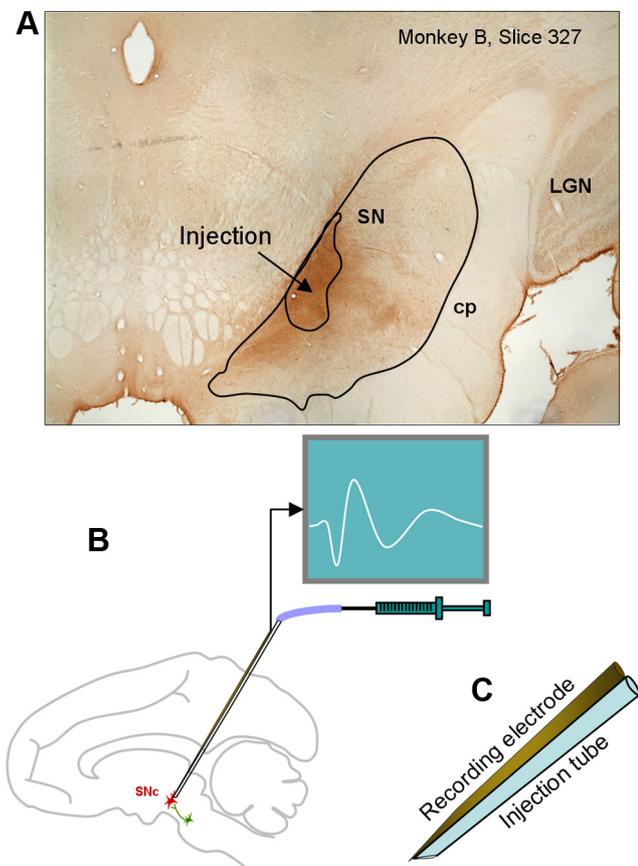


Figure 14. *A*, Injection site of CTB in right substantia nigra (SN). The injection site is indicated by an arrow in the figure. The injection was aimed at the location where a DA neuron (Fig. 10C) was recorded in one of the previous recording sessions where we verified bidirectional connectivity between the RMTg and SNc (Fig. 10E, F). *B*, To further insure the accuracy of the injection, we first recorded spike activity of a putative DA neuron before injecting CTB. *C*, This was possible because we used an injection tube that was attached to a recording electrode. Some spreading of the CTB was detected along the injection cannula, but it was deemed not to affect the results. LGN, Lateral geniculate nucleus; cp, cerebral peduncle.

Reward-positive neurons on the margins of the RMTg

Although our main focus was on reward-negative RMTg neurons that mediate the LHb-DA inhibition, reward-positive neurons were also found, usually away from the center of the RMTg (Fig. 13C). Unlike reward-negative RMTg neurons, which could be excited but were never inhibited by LHb stimulation, reward-positive neurons showed a mixture of excitations and inhibitions by LHb stimulation (Table 1). Several pieces of evidence suggest that these neurons are functionally distinct from the reward-negative neurons. First, some reward-positive neurons were inhibited by LHb stimulation. This inhibition is not likely due to direct LHb projections, which are glutamatergic, but due to indirect projections via an inhibitory intermediate (possibly the reward-negative RMTg neurons). Second, some reward-positive RMTg neurons were excited by LHb neurons, a paradoxical response given that excitations from the LHb should convey reward-negative responses, rather than the observed reward-positive signals. Hence, these reward-positive RMTg area neurons may receive additional inputs, such as excitatory inputs from reward-positive regions (possibly VP neurons), which would override the LHb inputs to produce reward-positive responses (Fig. 15). Third, some reward-positive RMTg neurons were activated antidromically from the SNc (Fig. 11). This does not fit the scheme of the RMTg→DA inhibition. A parsimonious

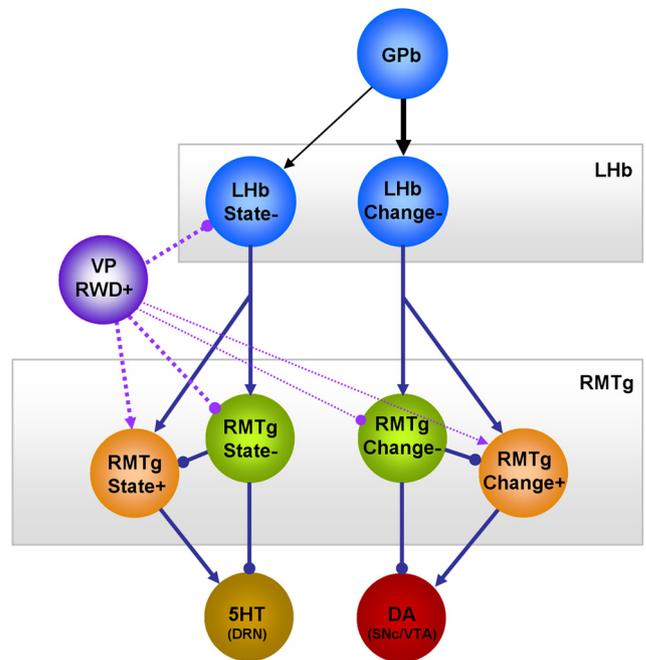


Figure 15. Speculative circuit diagram showing functional connectivity among subcortical motivation-related areas. “Change” indicates the change-of-value type, and “state” indicates the state-value type. For each of them, a minus sign indicates the reward-negative type and a plus sign indicates the reward-positive type. Excitatory connections are indicated by arrows; inhibitory connections are indicated by lines with filled circles. A dominant pathway is the border part of the globus pallidus (GPb)→Lhb change(−)→RMTg change(−)→DA, because the reward-negative neurons were more numerous than the reward-positive type in the RMTg. The reward-positive RMTg neurons showed mixed responses (Table 1) in reaction to the stimulation in the Lhb. This result can be explained if we assume two antagonizing connections to the reward-positive neuron: an excitatory input coming directly from the Lhb and an inhibitory input from the RMTg change(−) type. We speculate that the state value signals originate partly from the VP and Lhb. This is based on our unpublished observations that some neurons in the VP and some neurons in the more medial part of Lhb represent state values, and these areas are considered to project to the RMTg. We also speculate that the state value RMTg neurons modulate 5-HT neurons in the DRN, because neurons in the DRN encode state values, and the RMTg is known to project to the DRN. VP RWD+, Ventral pallidum reward-positive neurons.

explanation is that the reward-positive neurons projecting to the SNc/VTA may be excitatory, in line with a previous finding in the rat (Jhou et al., 2009b). In summary, reward-positive neurons tend to reside away from the RMTg center, tend to be less dominated by Lhb input, and may be excitatory rather than inhibitory to DA neurons.

Representation of state value in the RMTg

A subset of RMTg neurons showed sustained changes in activity along the different stages of a trial roughly corresponding to the levels (or states) of the expected reward values (Bromberg-Martin et al., 2010a). How can these RMTg neurons encode state value if the major inputs to the RMTg originate from the Lhb? There are several ways to explain this paradox (Fig. 15). First, some Lhb neurons might encode state value rather than change value, and they may project to the state-value RMTg neurons. Recently, we indeed found state-value neurons in the Lhb (S. Hong and O. Hikosaka, unpublished observation). Second, RMTg neurons might receive inputs from the VP (Jhou et al., 2009a; Kaufing et al., 2009), where neurons tend to encode state value.

We hypothesize that the target of these state-value RMTg neurons may be the dorsal raphe nucleus (DRN), which receives a strong direct projection from the RMTg (Kirouac et al., 2004;

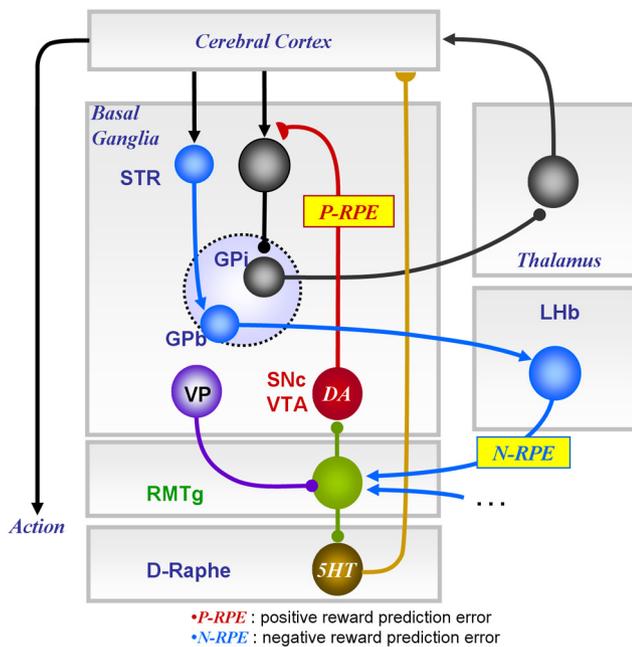


Figure 16. Speculative circuit diagram showing the functional connectivity in the basal ganglia and surrounding motivation-related areas. There are two functionally distinct motivation circuits, one for motor execution (matrix part of striatum→Gpi→the motor thalamus or brainstem nuclei) and the other for reward evaluation (patch part of striatum→border part of the globus pallidus (Gpb)→Lhb→RMTg→DA). Excitatory, inhibitory, and modulatory connections are illustrated with arrow heads, filled circles, and half circles, respectively. The ellipsis indicates many known inputs to the RMTg whose functions have yet to be examined. For simplicity, some known connections, such as the projection from the nucleus accumbens to the VP and the projections of the VP to DA neurons and to the Lhb, are omitted. STR, Striatum; N-RPE, negative reward prediction error; P-RPE, positive reward prediction error; D-Raphe, dorsal raphe.

Jhou et al., 2009a). Indeed, studies from our laboratory have shown that DRN neurons encode state value positively or negatively (Bromberg-Martin et al., 2010a), similarly to the state-value encoding RMTg neurons. In support of this hypothesis, it has been shown that the Lhb exerts a strong influence on serotonin release (Yang et al., 2008), and this is likely to be mediated by the connection from the Lhb to the DRN (Wang and Aghajanian, 1977; Herkenham and Nauta, 1979; Ferraro et al., 1996; Kim, 2009), which may be partly indirect, mediated by the RMTg (Kirouac et al., 2004; Jhou et al., 2009a). Hence, the RMTg may provide change-of-value signals to DA neurons in the SNc/VTA and state-value signals to serotonin neurons in the DRN (Figs. 15, 16), therefore influencing a majority of modulatory centers in the brain.

Hypothesis of reinforcement learning

Our data are supportive of a larger conceptual framework in which reward-related behavior involves two functionally distinct pathways, one involved in motor execution and the other involved in reward evaluation (Fig. 16). The motor execution pathway consists of the matrix part of striatum→globus pallidus internal segment (Gpi)→thalamus or brainstem connections (the pathway in black). The reward evaluation pathway consists of the patch (striosome) part of striatum (Rajakumar et al., 1993; Hong and Hikosaka, 2008b)→border part of the globus pallidus (Gpb) (Hong and Hikosaka, 2008a)→Lhb→RMTg→SNc/VTA (dopamine)→striatum connections. Along this extrabasal ganglia pathway, sensorimotor information is removed, and reward information is extracted at the Gpi→Lhb level (Hong and

Hikosaka, 2008a). The Lhb passes the reward evaluation signal to the RMTg. Upon getting the signal, the RMTg inhibits the DA using its large population of GABAergic neurons to reinforce or discourage the ongoing action via the dopamine projections to the striatum. It is also known that the RMTg projects to other neuromodulatory systems, such as the raphe nuclei and locus ceruleus (Jhou et al., 2009a; Kauffling et al., 2009). By providing unexpected reward results to these modulatory systems, the RMTg may regulate the mood of the animal (via serotonin system) and the level of attention (e.g., via noradrenalin) (Aston-Jones et al., 1999) to better adapt to changing environments.

References

- Aston-Jones G, Rajkowski J, Cohen J (1999) Role of locus coeruleus in attention and behavioral flexibility. *Biol Psychiatry* 46:1309–1320.
- Balcita-Pedicino JJ, Omelchenko N, Bell R, Sesack SR (2011) The inhibitory influence of the lateral habenula on midbrain dopamine cells: ultrastructural evidence for indirect mediation via the rostromedial mesopontine tegmental nucleus. *J Comp Neurol* 519:1143–1164.
- Belova MA, Paton JJ, Salzman CD (2008) Moment-to-moment tracking of state value in the amygdala. *J Neurosci* 28:10023–10030.
- Brinschwitz K, Dittgen A, Madai VI, Lommel R, Geisler S, Veh RW (2010) Glutamatergic axons from the lateral habenula mainly terminate on GABAergic neurons of the ventral midbrain. *Neuroscience* 168:463–476.
- Bromberg-Martin ES, Hikosaka O, Nakamura K (2010a) Coding of task reward value in the dorsal raphe nucleus. *J Neurosci* 30:6262–6272.
- Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O (2010b) Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* 67:499–510.
- Christoph GR, Leonzio RJ, Wilcox KS (1986) Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J Neurosci* 6:613–619.
- Davies RM, Gerstein GL, Baker SN (2006) Measurement of time-dependent changes in the irregularity of neural spiking. *J Neurophysiol* 96:906–918.
- Ferraro G, Montalbano ME, Sardo P, La Grutta V (1996) Lateral habenular influence on dorsal raphe neurons. *Brain Res Bull* 41:47–52.
- Geisler S, Trimble M (2008) The lateral habenula: no longer neglected. *CNS Spectr* 13:484–489.
- Herkenham M, Nauta WJ (1977) Afferent connections of the habenular nuclei in the rat. A horseradish peroxidase study, with a note on the fiber-of-passage problem. *J Comp Neurol* 173:123–146.
- Herkenham M, Nauta WJ (1979) Efferent connections of the habenular nuclei in the rat. *J Comp Neurol* 187:19–47.
- Hikosaka O, Nakamura K, Nakahara H (2006) Basal ganglia orient eyes to reward. *J Neurophysiol* 95:567–584.
- Hong S, Hikosaka O (2008a) The globus pallidus sends reward-related signals to the lateral habenula. *Neuron* 60:720–729.
- Hong S, Hikosaka O (2008b) Convergent inputs from the ventral striatum and the dorsal striatum to the lateral habenula in the monkey. *Soc Neurosci Abstr* 34:578.6.
- Jhou TC, Geisler S, Marinelli M, Degarmo BA, Zahm DS (2009a) The mesopontine rostromedial tegmental nucleus: A structure targeted by the lateral habenula that projects to the ventral tegmental area of Tsai and substantia nigra compacta. *J Comp Neurol* 513:566–596.
- Jhou TC, Fields HL, Baxter MG, Saper CB, Holland PC (2009b) The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* 61:786–800.
- Ji H, Shepard PD (2007) Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *J Neurosci* 27:6923–6930.
- Kauffling J, Veinante P, Pawlowski SA, Freund-Mercier MJ, Barrot M (2009) Afferents to the GABAergic tail of the ventral tegmental area in the rat. *J Comp Neurol* 513:597–621.
- Kim U (2009) Topographic commissural and descending projections of the habenula in the rat. *J Comp Neurol* 513:173–187.
- Kirouac GJ, Li S, Mabrouk G (2004) GABAergic projection from the ventral tegmental area and substantia nigra to the periaqueductal gray region and the dorsal raphe nucleus. *J Comp Neurol* 469:170–184.
- Lisoprawski A, Herve D, Blanc G, Glowinski J, Tassin JP (1980) Selective

- activation of the mesocortico-frontal dopaminergic neurons induced by lesion of the habenula in the rat. *Brain Res* 183:229–234.
- Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447:1111–1115.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Olson VG, Nestler EJ (2007) Topographical organization of GABAergic neurons within the ventral tegmental area of the rat. *Synapse* 61:87–95.
- Omelchenko N, Bell R, Sesack SR (2009) Lateral habenula projections to dopamine and GABA neurons in the rat ventral tegmental area. *Eur J Neurosci* 30:1239–1250.
- Paxinos G, Huang X, Toga A (1999) The rhesus monkey brain in stereotaxic coordinates, Ed 1. London: Academic.
- Perrotti LI, Bolanos CA, Choi KH, Russo SJ, Edwards S, Ulery PG, Wallace DL, Self DW, Nestler EJ, Barrot M (2005) DeltaFosB accumulates in a GABAergic cell population in the posterior tail of the ventral tegmental area after psychostimulant treatment. *Eur J Neurosci* 21:2817–2824.
- Rajakumar N, Elisevich K, Flumerfelt BA (1993) Compartmental origin of the striato-entopeduncular projection in the rat. *J Comp Neurol* 331:286–296.
- Tachibana Y, Hikosaka O (2009) The rostral part of the external pallidum and the ventral pallidum convey reward-associated visuomotor information. *Soc Neurosci Abstr* 35:567.564.
- Wang RY, Aghajanian GK (1977) Physiological evidence for habenula as major link between forebrain and midbrain raphe. *Science* 197:89–91.
- Yang LM, Hu B, Xia YH, Zhang BL, Zhao H (2008) Lateral habenula lesions improve the behavioral response in depressed rats via increasing the serotonin level in dorsal raphe nucleus. *Behav Brain Res* 188:84–90.