Synaptic Mechanisms for Auditory-Vocal Integration and the Correction of Vocal Errors

RICHARD MOONEY

Department of Neurobiology, Duke University Medical Center, Durham, North Carolina 27710, USA

ABSTRACT: A central goal of neuroscience is to understand the cellular mechanisms enabling the cultural transmission of behaviors, such as speech and music. Birdsong is a rare non-human instance of a culturally transmitted vocal behavior. The songbird’s brain provides a powerful system in which to study the cellular mechanisms underlying auditory-guided vocal learning. Identifying those mechanisms requires an analysis of synaptic function, because the synapse is the fundamental organizational unit of the neuronal networks that mediate behavior. Intracellular recordings provide a powerful method for simultaneously probing the activities of a single neuron and the synaptic networks in which that cell is embedded. This chapter details initial steps in the in vivo intracellular analysis of the synaptic connectivity of neurons important to singing and song learning. Our analysis is focused upon HVC and involves studies of interneurons as well as projection neurons of the two major output pathways of HVC. We test predictions derived from several models of how such learning may take place, including contributions from “comparator” and “corollary discharge” auditory feedback cancellation mechanisms. Our studies in anesthetized animals and brain slices provide insight into the synaptic properties of HVC that might be well suited for these mechanisms, although extrapolation to synaptic behavior in the awake, singing bird must be made with caution. We suggest that future work must extend the analysis of synaptic properties into the intact brain of the songbird, preferably as the bird learns to sing.

KEYWORDS: song system; synapse; postsynaptic potentials; intracellular recordings; excitatory; inhibitory; auditory feedback; sensorimotor integration; vocal learning; corollary discharge; cancellation

MODELS AND MECHANISMS OF SONG LEARNING

Songbirds learn to sing by using auditory feedback to match their own song to a memorized tutor model.1,2 Because many components of the brain circuits mediating this behavior have been identified, song acquisition provides a powerful model system for studying the neural mechanisms of auditory-guided vocal learning. Song learning in birds raises two central questions: First, what is the neural substrate of the memorized tutor model (i.e., the acquired template, see Adret, this volume). Sec-
ond, how does the developing bird’s brain use auditory feedback to acquire and maintain those features of the species-specific song transmitted by the tutor, a process which requires it to control its own vocal output during development and as an adult (see Woolley and Konishi, this volume). In this chapter we describe a number of studies designed to test, at the synaptic level, several possible models of the mechanisms mediating song learning and maintenance.

One such model is that of a comparator circuit that uses auditory information generated during production of the bird’s own song (BOS) to detect differences between the currently emitted song and the stored model. These differences generate an adap-

![Figure 1](image.png)

**FIGURE 1.** Models are shown of comparator and cancellation circuits that could underlie vocal learning through auditory feedback. (A) A hypothetical comparator circuit evaluates auditory feedback about the bird’s own song in the context of a memorized model (or template). Differences between the actual and desired output result in an error signal that adaptively modifies the vocal system so that its output better matches the model. (B) Cancellation of anticipated auditory feedback could be achieved by harnessing corollary discharge from song motor circuits (the “vocal system”). The corollary discharge would ultimately inhibit auditory activity evoked by the anticipated vocal output, thus minimizing auditory desensitization and maximizing auditory sensitivity to vocal errors.
tive signal that drives the song central pattern generator (CPG) towards the memorized tutor song (Fig. 1A). Two other related problems that the songbird’s brain is faced with during singing is how to discriminate self-generated sounds from those emitted by external sources, and how to prevent the auditory system from becoming desensitized to self-generated sounds during repeated bouts of singing. A neural mechanism that might play a role in solving these two problems uses corollary discharge from the song CPG to cancel out anticipated sensory feedback. This cancellation process would increase sensitivity to signals differing from the BOS, including errant self-generated vocalizations, thereby simplifying the detection of deviations from the template by the “comparator circuit.” Indeed, a cancellation mechanism like this plays an important role in the communication behavior of weakly electric fish, and in the maintenance of auditory sensitivity during “singing” in crickets and during vocalization in mammalian species, including bats and primates. Interestingly, chronic recordings made in the song nucleus HVC also hint at cancellation of auditory responsiveness during singing.

Each of these models generates certain predictions about the kinds of properties that their neural substrates might be expected to exhibit. For example, neurons in a comparator circuit might be expected to exhibit relatively broad responsiveness to auditory inputs in order to detect a wide range of vocal performances, many of which will deviate substantially from the target—the stored model. Similarly, if a cancellation mechanism is involved in song acquisition, we might expect that the neurons involved in this process will exhibit some evidence of a cancellation signal, perhaps in the form of synaptic inhibition yoked to the premotor command signal (Fig. 1B). In the expertly trained adult songbird, these cancellation signals would be optimally weighted to suppress auditory feedback arising from a correct rendition of the bird’s song, only letting vocal “errors” through to modify the adaptive signal generated by the comparator. Ultimately, the cancellation signal would be subtracted from the actual auditory feedback to simplify error detection and minimize auditory desensitization in sensorimotor areas important to the error correction process.

Tests of these hypotheses must involve analyses of synaptic activity at critical points within the “song circuit.” Although song learning is likely to be a distributed function within the songbird brain, the telencephalic nucleus HVC is a key place to start looking for the synaptic mechanisms governing auditory-vocal integration, song premotor-triggered auditory cancellation, and activation of error correction pathways. HVC contains two functionally distinct types of projection neurons: one class (HVCRA) innervates the forebrain premotor nucleus RA, which in turn gives rise to descending projections onto the brainstem respiratory-vocal network. During singing, HVCRA neurons generate highly precise bursts of premotor activity that are thought to be essential to song patterning (see Fee and colleagues, this volume). Another class of (HVCX) innervates a basal ganglia homologue (area X) within the AFP (see Brainard and Perkel, this volume). The output of the AFP is LMAN, whose axons innervate the same RA premotor neurons receiving direct HVC input, providing a potential cellular substrate for AFP modulation of vocal plasticity. HVC itself receives auditory and possibly proprioceptive input, the sources of which are only partially identified (see Wild and Goller & Cooper, this volume). Inputs to HVC are likely to undergo extensive local processing, because the nucleus contains a variety of interneuron types and axons from both projection neuron types extend collaterals within the nucleus. Thus, HVC sits at the apex of an auditory-vocal
pathway for song, provides the synaptic organization permissive for extensive local processing, and gives rise to pathways important for song patterning and for error correction. Therefore, HVC is an obvious place to begin a search for synaptic mechanisms important to auditory-vocal interactions important to song learning, including error correction processes.

The highly selective auditory response properties of HVC neurons constitute a striking example of experience-dependent sensory tuning. Both types of HVC projection neuron respond strongly and highly selectively to playback of the BOS. This makes HVC the likely source of the auditory responses detected in both the posterior and anterior pathways of the song circuit, an idea that still awaits complete experimental validation. The auditory information transmitted by HVC projection neurons is also highly specific: HVC neurons are “BOS-selective,” firing strongly to forward playback of the BOS, but not to temporally manipulated versions of the BOS, such as reverse song or reverse-syllable song, nor to other conspecific songs or synthetic sounds. Moreover, BOS selectivity in HVC is important to any discussion of auditory vocal integration, because of HVC’s obligatory role in song patterning. An intracellular analysis of auditory-evoked activity in identified HVC neurons is likely to unearth useful clues about the functional properties of the HVC circuit and to illuminate potential cellular sites of auditory-vocal integration.

One of the primary findings of this analysis is that HVCRA neurons respond to BOS playback, suggesting that the same premotor neurons that pattern song also respond to the auditory feedback resulting from their premotor activity. This is an attractive idea, because it places the site of auditory feedback (and thus perhaps of error correction) directly on the song patterning side of the circuit (i.e., the posterior pathway). However, instead of displaying the broad responsiveness to be expected in components of an error detecting (comparator) mechanism, HVCRA neurons discharge action potentials almost exclusively to playback of forward BOS, and not to its temporal variants. At first glance, such response exclusivity would seem incompatible with the need to detect errors in vocalization. On the other hand, the exclusivity manifested in the extracellular response could reflect the outcome of several putative underlying synaptic processes, at least some of which would be compatible with the requirement of broadly tuned input to a hypothetical comparator (Fig. 2A).

One such process posits auditory afferents arising outside of HVC, which though BOS-selective also synaptically transmit information to HVCRA neurons about a wide range of auditory stimuli, but at a subthreshold level. A variant of this process is a combination of non-selective extrinsic afferents to HVCRA that are modulated locally by HVC neurons to add an excitatory bias to the BOS or an inhibitory bias to non-BOS stimuli, resulting in exclusive firing to the BOS. Finally, one may conceive of a “pipeline” mechanism involving auditory inputs that fire action potentials as selectively as the HVCRA neurons they innervate. This last synaptic process is not compatible with the comparator input scheme, because HVCRA neurons would receive narrowly tuned auditory information that would not report deviances in the vocal performance. If HVC receives a pipeline of BOS-selective information, then other areas presynaptic to HVC might be more likely places to search for the auditory input to a comparator circuit.

To determine which of these three processes applies to HVCRA, we used intracellular recordings to examine their patterns of auditory-evoked synaptic activity. We found that HVCRA neurons fire a very phasic burst of action potentials almost exclu-
FIGURE 2. Hypothetical and observed synaptic processes underlying the genesis of narrowly tuned, BOS-selective spiking in HVCRA neurons. (A) Several models that could explain BOS-selective spiking in HVC include equally selective auditory afferents (left-“pipeline” model); afferents that are biased to the BOS, but that fire to other stimuli (middle); and unbiased afferents (right). In the two latter cases, additional processes, such as postsynaptic thresholding or other synaptic interactions local to HVC result in more exclusive BOS-evoked spiking. In each case, action potential activity is shown for the auditory afferent and sub- and suprathreshold activity is shown for an HVCRA neuron in response to both forward and reverse playback of the BOS. These two stimuli, which contain the same spectral energy but differ in their temporal organization, can be used to measure selectivity of HVC neurons.17,19 (B) The actual synaptic events underlying BOS-evoked spiking in HVCRA neurons are consistent with the middle of the three models shown in (A). Only the BOS and to a lesser degree reverse syllable order BOS (BOS-RS) evoke action potentials in this neuron (top row: PSTHs of suprathreshold activity). However, other stimuli, including conspecific (CON) and heterospecific (HET) birds’ songs and reversed BOS (BOSrev), evoke depolarizing and entirely subthreshold synaptic activity (middle row: median-filtered average membrane potential records). Therefore, auditory afferents to these neurons fire action potentials to these non-BOS stimuli, but display a bias to the BOS over the other stimuli. Oscillograms of the various songs are shown in the bottom row; neuronal responses are to 20 iterations of each stimulus.
MOONEY: SYNAPTIC MECHANISMS

sively to BOS playback, and then often only once per motif. This pattern is strikingly similar to the discharge pattern of this same cell type recorded in the singing bird (Fig. 2B and see figure in Fee and colleagues, this volume). It is possible that the apparent equivalence of BOS-related auditory responses and song premotor activity reflects a common coding scheme that functions to simplify the process of auditory-vocal integration. Additional studies are needed to establish whether there is a similar equivalence in the song-related auditory and motor activity of individual HVCRA neurons. We also found that, throughout much of their period of BOS-evoked discharge, HVCRA neurons undergo a sustained subthreshold depolarization, punctuated intermittently by the much stronger depolarizations associated with the cell’s phasic spiking (Fig. 2). Therefore, synaptic inputs to HVCRA neurons are activated throughout the BOS, although much of this activity remains subthreshold. Furthermore, in vivo intracellular recordings also revealed that a wide range of non-BOS auditory stimuli (including temporally manipulated versions of the BOS, other conspecific and heterospecific songs, and noise bursts) is effective in activating synapses on HVCRA neurons (Fig. 2B). Interestingly, almost all of the responses evoked by non-BOS stimuli remain subthreshold. That is, while the extracellular response appears BOS-“exclusive,” the underlying subthreshold synaptic activity—which can only be detected with an intracellular electrode—is relatively broadly tuned. These broad synaptic tuning properties potentially endow HVCRA neurons with the capacity to detect auditory information about vocal performances that deviate from the ideal song phenotype (i.e., the BOS)—one of the features associated with the input to our hypothetical comparator. More generally, the sharp contrast between narrow suprathreshold tuning and the much broader subthreshold tuning we have seen in HVCRA neurons is also a characteristic of higher order sensory processing in other systems, including the mammalian primary visual cortex and the owl auditory midbrain.

The broad synaptic response patterns of HVCRA neurons also may have implications for the type of auditory information that HVC can transmit to the song motor nucleus RA. Although much of the response to non-BOS stimuli remains entirely subthreshold in anesthetized animals, artificially generating a small positive offset in the cell’s resting membrane potential (i.e., by injecting positive current through the recording electrode) is sufficient to enable the cell to spike to non-BOS stimuli. This membrane potential effect indicates that the subthreshold response patterns of HVCRA neurons are largely due to excitatory synaptic inputs. Therefore, endogenous factors that shift the resting membrane potential of HVCRA neurons likely will influence whether these cells fire action potentials to auditory stimuli and thus transmit auditory information to RA. Consistent with this idea, Margoliash and his co-workers have noted a state-dependent gating of auditory responses in RA, with responses readily evoked in the sleeping bird but largely suppressed in the waking animal. Furthermore, other studies have shown that gating of auditory responses also occurs at the level of HVC, at least in the zebra finch. Perhaps the gating of RA’s auditory responses is due to shifts in the resting membrane potential of HVCRA neurons, with auditory signals propagating to RA when the HVCRA neuron membrane potential is shifted to a more positive and thus permissive state. The specific factors that trigger such modulation are unknown, but are likely to involve classic neuromodulators, such as acetylcholine and norepinephrine, both of which are abundant in HVC. Two key states that could augment the resting tone of HVCRA neurons in
the bird when it is awake are attention and song motor activity. In the future, it will be important to identify the endogenous molecules and mechanisms that mediate such modulation and to determine whether modulation of HVC<sub>RA</sub> occurs when juveniles learn to sing or when adults attend to their own songs or to other singing birds.

**ORIGINS OF SELECTIVE AUDITORY RESPONSIVENESS IN HVC<sub>RA</sub> NEURONS**

The broad subthreshold responsiveness of HVC<sub>RA</sub> neurons raises questions as to the source and selectivity of their auditory afferents. One possible arrangement involves a population of broadly tuned afferents arising extrinsic to HVC. A second involves the convergence upon HVC<sub>RA</sub> cells of presynaptic inputs (extrinsic or local) from multiple neurons, each of which is narrowly tuned, but to different acoustical features of the BOS and/or other stimuli. One method for distinguishing extrinsic from local contributions to song-evoked synaptic activity would be to inactivate the HVC local circuit without silencing the activity of extrinsic auditory afferents. In the inactivated HVC, intracellular recordings could be used to measure the contributions of the extrinsic inputs, as reflected by any persistent subthreshold au-

**FIGURE 3.** A pharmacological method for distinguishing local and extrinsic contributions to auditory responses in HVC is shown. Song-evoked synaptic activity in an HVC projection neuron is likely to reflect contributions from both extrinsic and local (intranuclear) components (left). Local contributions to the net synaptic activity include synapses from interneurons, and the local collaterals of other projection neurons; extrinsic contributions are from HVC’s auditory afferents. Application of a concentrated GABA solution to HVC through a puffer pipette (right) can be used to silence local activity. Remaining auditory-evoked activity then can be attributed to the extrinsic inputs onto the impaled HVC neuron. The song stimulus is depicted as an oscillogram at the bottom of the figure.
MOONEY: SYNAPTIC MECHANISMS

The comparison between song-evoked response patterns in the normal and inactivated state could also be used to infer how the local circuit contributes to the generation of BOS selectivity. To inactivate the circuit, we recorded intracellularly from identified HVC neurons, then pressure-ejected a small amount of concentrated GABA into HVC (Fig. 3). This treatment reversibly shunts the dendritic and somatic membranes of the HVC cells that the GABA contacts, preventing synaptically evoked spiking. In this state, local processing is disabled, but axons arising from outside of HVC, the terminals of which presumably lack GABA receptors, should continue to function. If local convergence of narrowly tuned inputs accounts for the broad subthreshold responses of HVC_RA neurons, then subthreshold responses should become more narrowly tuned (i.e., show more phasic synaptic responses to the BOS and respond to a reduced set of non-BOS stimuli) upon inactivation. In contrast, if HVC_RA subthreshold response patterns are dictated by extrinsic afferents, they should persist unchanged. In fact, the basic subthreshold response pattern seen in HVC_RA neurons prior to GABA application persisted during inactivation of the HVC local circuit, indicating to us that the extrinsic inputs into HVC are BOS selective, but also fire action potentials to a wider range of song and non-song stimuli. Therefore, HVC’s auditory afferents as a population are broadly responsive to different auditory stimuli, rather than being narrowly tuned. Only single unit recordings can resolve the remaining question of whether individual afferents to HVC are broadly or narrowly tuned.

NIF CONTRIBUTIONS TO BOS SELECTIVITY IN HVC

The finding that auditory afferents to HVC are BOS selective is intriguing because prior studies suggested that heightened selectivity for the BOS originated within HVC. Importantly, the pronounced BOS selectivity seen in HVC appears to be largely absent in the primary auditory telencephalic region, Field L, which is presumed to be a major source (direct or indirect) of auditory input to HVC. Recent studies have suggested nucleus interfacialis (NIf) as the site where the auditory signal acquires BOS selectivity. NIf, which has been identified in tracing studies as a major HVC afferent, contains neurons that display BOS selectivity intermediate between that of Field L and HVC neurons. These features of NIf suggested that it might be involved in the transformation of non-selective input from Field L into a moderately BOS-selective output that is further augmented by the HVC local circuit. Another possibility is that HVC receives auditory input from NIf and several other brain areas, and together these various auditory inputs give rise to the synaptic responses observed in HVC_RA neurons. In support of this latter view, at least one other HVC afferent, the nucleus mMAN, is known to contain neurons that are BOS selective.

To determine the extent to which NIf could account for the synaptic activity of HVC_RA neurons and to assess any changes in auditory selectivity between these two song nuclei, we paired extracellular recordings in NIf, to measure its action potential output, with intracellular recordings from HVC_RA neurons, to measure their synaptic input. These recordings revealed a striking concordance in the auditory properties measured at these two sites with respect to both the temporal pattern and relative strength of their auditory-evoked activities. For example, a given stimulus evoked a
mean firing rate in NIf that closely paralleled the strength of the synaptic response in HVC. In addition, auditory-evoked responses in both areas were graded, with the BOS evoking the highest levels of activity, and reverse-syllable BOS, reverse BOS, and conspecific songs eliciting weaker responses. Furthermore, NIf spikes preceded synaptic activity in HVC by one to several milliseconds, and inactivating NIf with concentrated GABA abolished all spontaneous and auditory-evoked synaptic and spiking activity in HVC, suggesting to us that NIf is the dominant and perhaps sole source of ascending auditory input to HVC. Indeed, these dual recordings show that

**FIGURE 4.** Two models depicting the generation of sparse spiking patterns in HVC<sub>RA</sub> neurons in response to BOS playback. (A) Several neurons presynaptic to the HVC<sub>RA</sub> neuron spike in a continuous fashion in response to the stimulus, but only rarely fire at the same time. The HVC<sub>RA</sub> neuron only fires action potentials when it detects this coincident activity. (B) Another possibility is that sustained firing in the presynaptic only intermittently shifts the HVC<sub>RA</sub> neuron membrane potential above spike threshold, perhaps through temporal integration by the postsynaptic cell. In both examples, the song stimulus is depicted as an oscillogram at the bottom of the panel.
HVC\textsubscript{RA} neurons act largely as followers of BOS-selective input from NIf, in contrast to the idea that BOS selectivity arises or is further enhanced in HVC. One implication of these results is that the search for the origins of BOS selectivity must now extend to NIf’s auditory inputs, which are reported to include Field L, cHV, and perhaps even direct projections from the auditory thalamus\textsuperscript{36} (Wild, personal communication).

While the relative bias to the BOS does not appear to be enhanced at the NIF-HVC\textsubscript{RA} synapse, this synapse does appear to be the site where an almost exclusive action potential response to self-generated vocalizations first emerges. That is, whereas NIf neurons fire to a wide range of auditory stimuli, HVC\textsubscript{RA} neurons fire almost exclusively to the BOS. The mechanisms underlying this transformation in stimulus specificity remain unknown. One possibility is that subsets of NIf neurons converge on individual HVC\textsubscript{RA} neurons, but only drive spiking when their activities are highly correlated, perhaps because HVC\textsubscript{RA} neurons act as coincidence detectors (FIG. 4A). Another hypothesis is that the resting membrane potential of HVC\textsubscript{RA} neurons and/or the strength of NIf synapses is adjusted so that only the BOS generates sufficient depolarization to exceed spike threshold (FIG. 4B). The observation that BOS-evoked spiking in an HVC\textsubscript{RA} neuron becomes sustained throughout more of the stimulus when the cell’s resting potential undergoes a positive DC offset (via current injection) is supportive of the thresholding model and contradicts a model where HVC spike timing is determined entirely by presynaptic cooperativity.\textsuperscript{14} In either case, the sustained synaptic activity evoked by BOS playback is transformed into a much more intermittent, or temporally sparser, action potential output by the HVC\textsubscript{RA} neuron. An added motivation to understanding how sparse spiking is generated in response to BOS playback is that a similar process likely accounts for the extremely sparse spike patterns HVC\textsubscript{RA} neurons produce during singing.\textsuperscript{23}

**AUDITORY SELECTIVITY AND CANCELLATION MECHANISMS IN THE HVC\textsubscript{X} (ANTERIOR) PATHWAY**

Intracellular analyses of synaptic mechanisms in HVC\textsubscript{X} neurons will be critical in the study of song learning. These neurons are the major and perhaps sole source of auditory input to the AFP, which is crucial to the vocal plasticity necessary to song learning.\textsuperscript{37,38} That HVC\textsubscript{X} neurons may make different but complementary contributions from HVC\textsubscript{RA} neurons to the song learning process is suggested by the observation that though BOS playback evokes very phasic action potential responses from both projection neurons, HVC\textsubscript{X} cells are quite distinct in their song-evoked synaptic activity from HVC\textsubscript{RA} neurons. Specifically, HVC\textsubscript{X} neurons exhibit a hyperpolarizing response in membrane potential that is sustained throughout much of the BOS, suggestive of a synaptic inhibitory process.\textsuperscript{14,39} Such inhibition could contribute to a cancellation component of a comparator model under two conditions. First, the inhibition should be most strongly evoked by the BOS, rather than other auditory stimuli; second, the inhibition should be activated by song premotor (HVC\textsubscript{RA}) neurons.

Consistent with the first condition, we found that the auditory-evoked inhibition detected as a postsynaptic hyperpolarization in HVC\textsubscript{X} cells is greatest for the BOS, by comparison with non-preferred stimuli such as reverse BOS, other songs, or noise bursts.\textsuperscript{14,39} In fact, the only stimulus other than the BOS that consistently evokes any
hyperpolarizing response in HVCX neurons is reverse syllable order BOS, which preserves the local temporal organization of song syllables, but does not maintain their overall sequence. The capacity of the reverse-syllable order BOS to evoke substantial inhibition suggests that the neurons driving the inhibition are tuned in part to the structure of individual syllables. Such a syllable-tuned inhibitory network could still operate to cancel out syllables that were correctly produced, even if they were not generated in the correct sequence. These findings are not readily consistent with a model where inhibition actively cancels out excitatory responses to the non-preferred stimuli, which is one way that inhibition could enhance BOS selectivity in HVC.

A second corollary prediction of a cancellation model is that inactivating the inhibition onto HVCX neurons should unmask an even stronger BOS-evoked synaptic excitation. We investigated this possibility using concentrated GABA to inactivate the entire HVC local network, while recording intracellularly from HVCX neurons.31,40 This treatment completely abolished BOS-evoked hyperpolarizations in HVCX neurons, suggesting that they originated from local circuit activity. The finding of a strong correlation between interneuron firing rates and HVCX membrane hyperpolarizations supported this conclusion.14 In fact, as predicted above, inactivating the local circuit did more than merely abolish the hyperpolarizing response; it unmasked BOS-evoked depolarizations in HVCX neurons that were sustained throughout the stimulus presentation. These stronger, more sustained depolarizing responses were remarkably like those we recorded from HVCRA neurons in the same animal, suggesting that the two projection neuron types receive common excitatory input, but that HVCX cells also receive an additional inhibitory component. This finding challenges the assumption that HVCX neurons only receive phasic excitatory synaptic activity. It suggests, instead, that the few excitatory peaks seen in control recordings are part of a more sustained synaptic excitation that is largely suppressed by inhibition.

A final important prediction of the cancellation hypothesis is that removing inhibition should increase BOS-evoked firing relatively more than firing evoked by other stimuli. That is, in the absence of inhibition, HVCX neurons should become even more BOS selective. The results of the local inactivation experiments (involving GABA) are supportive but remain inconclusive, because GABA treatment abolished all spiking. This prevented us from testing the prediction by measuring BOS-evoked firing in HVCX neurons. To address this question we need to selectively block inhibition onto HVCX neurons while allowing them to continue firing action potentials. Furthermore, to avoid non-specific network effects that may accompany pharmacological disinhibition, it would be important to apply intracellular blockers to single cells rather than to the nucleus. Fortunately, this technically challenging approach was made easier by a series of in vitro studies that identified three major classes of inhibitory input onto HVCX neurons.12,41,42 The first of these inhibitory inputs is an ionotropic GABA_A receptor that mediates a chloride current, while the other two are metabotropic receptors that bind glutamate (mGluR) or GABA (GABAB) to activate a G-protein–coupled inward rectifying potassium current (i.e., a GIRK). To determine which of these inhibitory processes mediates the auditory-evoked hyperpolarizing inhibition seen in HVCX neurons, we introduced a series of compounds through the recording pipette that selectively blocked either chloride or potassium channels, or disrupted G-protein–mediated signaling. These studies demonstrated
FIGURE 5. Intracellular disruption of G-protein–mediated potassium currents in HVC_X neurons augments their excitatory responses to BOS playback. (A) Occlusion of G-protein coupled inward rectifying potassium currents (GIRKs) with GTPγS abolishes BOS-evoked hyperpolarizations in HVC_X neurons and increases the strength of their BOS-evoked firing. The action potential PSTH (top) and median-filtered average membrane potential records (bottom) are shown either immediately after impaling the cell (0-5 min), and then again after the GIRK-mediated inhibition was occluded (12-17 min). Positive current was applied through the recording electrode to maintain the resting membrane potential at -63 mV.
that GIRKs are likely to be the major source of auditory-evoked hyperpolarization in HVCx neurons.\cite{Rosen:2008} When HVC\textsubscript{X} neurons were treated with compounds that disrupt GIRK activation, the strong BOS-evoked hyperpolarizing responses rapidly waned, unmasking a sustained depolarizing response like that seen with the local circuit inactivated. Moreover, BOS-evoked firing activity of the cell was strongly and preferentially enhanced, so that disruption of the inhibitory process actually increased the suprathreshold selectivity for the BOS (FIG. 5). This observation demonstrates that the GIRK-mediated inhibition actively dampens or cancels the BOS-evoked response, a finding consistent with a cancellation process. [Note: We found that GABA\textsubscript{A}-mediated chloride components are also activated by BOS playback, but do not hyperpolarize the cell, perhaps because the chloride reversal potential in HVC\textsubscript{X} neurons is positive of the resting potential.]

THE ROLE OF SONG PREMOTOR ACTIVITY IN THE CANCELLATION PROCESS

As noted earlier, such tuned inhibition, if effectively recruited by song premotor activity, could cancel out anticipated auditory feedback in a manner consistent with the comparator model. We therefore sought evidence that the inhibition that suppresses BOS-evoked firing in HVC\textsubscript{X} neurons is driven by song premotor neurons (i.e., HVC\textsubscript{RA}) neurons. Suggestive evidence for such an interaction between the two neuronal populations comes from simultaneous recordings from the two projection neurons types. These recordings show that BOS playback evokes opposing trajectories in their membrane potentials, almost as if the synaptic patterns in the two cell types were mirror images of each other (FIG. 6).\cite{Rosen:2008} That is, when HVC\textsubscript{RA} neurons are excited, HVC\textsubscript{X} neurons are inhibited, and vice versa, resulting in alternating BOS-evoked firing patterns. Two kinds of evidence suggest that the inhibition of HVC\textsubscript{X} neurons is driven by excitatory signals from HVC\textsubscript{RA} neurons inverted through an inhibitory interneuron network. First, HVC\textsubscript{RA} neurons make excitatory synapses in RA\cite{Rosen:2008} and on HVC interneurons (unpublished observations). Second, interneuron firing rates positively correlate with HVC\textsubscript{X} cell membrane potential negativity.\cite{Rosen:2008}
To obtain a more detailed view of the HVC microcircuit generating these synaptic interactions, we made intracellular recordings from identified HVC neurons in brain slices. We identified the synaptic connections of HVC RA neurons upon other HVC neurons by electrically stimulating the fiber tract that connects HVC RA neurons to RA. This stimulation evokes antidromic action potentials that invade the local collaterals of HVC RA axons, which then synaptically activate other HVC neurons. We found that such stimulation evokes short latency excitatory postsynaptic potentials (EPSPs) in HVC interneurons, and longer latency inhibitory PSPs (IPSPs) in HVC X neurons. These longer latency IPSPs were abolished by blockers of ionotropic glutamate receptors, suggesting to us that HVC RA neurons drive feedforward inhibition in HVC X neurons through synaptically interposed interneurons. To more directly test this idea, we made paired recordings from different HVC cell types. These recordings confirmed that HVC RA neurons make excitatory synapses onto interneurons, and that interneurons make inhibitory synapses onto HVC X neurons. Interestingly, the IPSPs that interneurons evoke in HVC X neurons are blocked by GABA_A receptor antagonists and have onset latencies too short to be mediated by second

FIGURE 6. BOS playback evokes different synaptic processes and reciprocal suprathreshold activity in HVC projection neurons. The two different HVC projection neuron types fire in alternating fashion to song playback (top: PSTH; the BOS is shown as an oscillogram at the bottom of the figure), while HVC_RA neurons undergo sustained depolarizing synaptic activity and HVC_X neurons display sustained membrane hyperpolarizations interrupted by phasic depolarizations (middle: median filtered average membrane potential records). The membrane potential hyperpolarizations in HVC_X neurons closely correlate with the peaks of maximum firing activity in HVC interneurons (bottom: PSTH), which in turn correlates closely with firing in HVC_RA neurons. These records are consistent with the idea that HVC_RA neurons drive feedforward inhibition in HVC_X neurons through an intervening interneuron layer. Responses shown are to 20 iterations of song playback. (Reprinted from Mooney with permission.)
messengers. Thus they are not the source of the GIRK-mediated inhibition evoked by song playback. One hint as to its possible source comes from the observation that, in the presence of GABA A receptor blockers, slow IPSPs are evoked in HVC X neurons by antidromic stimulation of HVC RA neurons. We do not yet know the pharmacology of these slow IPSPs, but they do not seem to be due to direct coupling between HVC RA and HVC X neurons (i.e., via metabotropic glutamate receptors), because ionotropic glutamate receptor antagonists block them (unpublished observations). One possibility is that another class of interneuron is excited by HVC RA neurons and activates either mGluR- or GABA B receptors on HVC X neurons. Another possibility is that the high instantaneous firing frequencies interneurons achieve in vivo result in excess transmitter release (i.e., synaptic spillover) that activates GABA receptors in addition to the GABA A variety.

In summary, we have found several synaptic features in HVC that are consistent with a hypothetical comparator model, including broadly tuned auditory input and inhibitory mechanisms that appear well suited for canceling anticipated auditory feedback. Inactivation experiments reveal that both HVC RA and HVC X neurons are innervated by extrinsic auditory inputs that respond broadly to many different auditory stimuli, not just the BOS. Broadly tuned auditory excitatory drive to the two HVC PNs likely arises from NIf, and can convey auditory information about errant as well as “correct” versions of the BOS. The synaptic architecture of HVC includes multiple forms of inhibition onto HVC X neurons, and inhibition can be recruited by HVC RA neurons, raising the possibility that singing-related activity in these HVC RA neurons will be translated into inhibitory signals in HVC X neurons. This process also could provide part of the synaptic basis for the corollary discharge that is seen in the AFP during singing. Finally, second-messenger–mediated inhibition in HVC X neurons suppresses BOS-evoked excitation, rather than acting to enhance it, a behavior reminiscent of cancellation processes seen in other sensorimotor pathways.

PROBLEMS AND FUTURE DIRECTIONS

Despite the obvious progress represented by these findings, it is important to recognize that our current understanding of synaptic processes in HVC derives entirely from either brain slices or anesthetized whole animal preparations, and not from intracellular recordings in the singing bird. Thus it is still unknown whether song playback in the anesthetized birds activates only the auditory feedback arm of the system or instead reactivates HVC in a manner that more fully mimics its activity during singing. The intracellular recordings in singing birds that could help answer this question are likely to remain beyond our technical capacities for the near future. However, intracellular or optical recording methods may eventually provide us with insights into the subthreshold activity of HVC neurons in the singing animal. In the interim, extracellular recordings from identified neurons should help to clarify the relation of real time auditory feedback to HVC and cancellation processes.

Several predictions of the comparator model could be tested using extracellular recordings from identified HVC neurons in awake songbirds. One prediction of the cancellation model is that distorted auditory feedback and/or vocal errors induced by peripheral manipulations (i.e., tracheosyringeal nerve section or injecting air into the respiratory system) should alter the activity of HVC X neurons. Another potentially
testable prediction is that interfering with GIRK signaling selectively in HVC$_X$ neurons should destabilize vocal output, because cancellation of auditory feedback would be disrupted. A third prediction is that natural states that tonically depolarize HVC$_{RA}$ neurons should gate auditory activity to RA. While the most obvious of these states is likely to be singing, it is also possible that attention serves to modulate the resting tone of these neurons, perhaps when the bird is in a practice rather than a performance state.

Another outstanding issue is the relative importance of online auditory feedback to song learning and song maintenance. In the present model, we envision that pre-motor-linked inhibition in HVC$_X$ neurons would be used in real time to cancel out auditory feedback. As discussed in Konishi’s article in this volume, the real-time auditory feedback model is problematic at moderate to high rates of syllable generation, because auditory feedback about a given vocal element coincides with the current motor pattern, not the earlier motor pattern that produced the vocal element itself. One way around this problem is to create short-term synaptic “memories” of the motor signal that flag which synapses are to be acted upon by the feedback signal. “Tagging” of previously active synapses has been proposed to underlie synapse specificity of certain forms of hippocampal long-term potentiation. If a similar process operates in HVC to correctly assign auditory feedback to previously active synapses it likely would require extraordinarily precise temporal regulation of these tags. Another possibility is that HVC’s inhibitory pathways are delay-tuned to match the corollary discharge to the appropriate auditory feedback. In this scenario, HVC$_X$ neurons would integrate auditory feedback with the delayed and inverted representation of the song premotor signal emanating through HVC interneurons. This appears to be similar to the cancellation process, but with the outcome that the balance of motor-linked inhibition and auditory-evoked excitation in HVC$_X$ neurons would constitute an estimate of the error signal, not simply the removal of a bias signal. This does not solve the central problem of how such an error signal then modifies the motor network, however. In addition, any model will need to incorporate the synaptic transfer function of the AFP to better understand the significance of inhibition and excitation in HVC$_X$ neurons. By analogy to the mammalian cortical-basal ganglia pathways, HVC$_X$ neurons may trigger LMAN activity through disinhibitory mechanisms. In this case, the pattern of synaptic inhibition in HVC$_X$ neurons may have special consequences for understanding the activation of circuits important to audition-dependent vocal plasticity.

A general issue that also remains unresolved is the exact pathway via which auditory information enters the song system. Our present studies suggest that NIf serves as the dominant and perhaps sole source of ascending auditory input to HVC. These studies do not address the full extent of other, perhaps weaker, ascending auditory inputs to HVC, including those from Field L and the shelf, and they also cannot address the contribution made by recurrent pathways (i.e., from mMAN). Still unresolved is the means via which auditory information enters NIf. Clearly, the lack of detailed understanding of ascending auditory projections to song nuclei stands in stark contrast to the increasingly detailed picture that is emerging of synaptic connectivity within and between the song nuclei.

Perhaps the most important aspect of the comparator circuit not touched upon here is the physical representation of the memorized model to which the BOS is being compared (see Adret, this volume). If auditory inputs to HVC convey informa-
tion only about the current song as part of a comparator process, it is reasonable to assume that the representation of the model must be stored either in or beyond (i.e., postsynaptic) to HVC. Furthermore, if NIf only conveys information about the current song, then perhaps other auditory inputs to HVC, which may include nMAN, Field L, and perhaps nucleus ovoidalis (see Wild, this volume), could convey a representation of the song model to HVC. In this view, HVC would receive two distinct inputs, the first conveying information about the model, and the second conveying information about the current BOS. A further extension of the view is that these two inputs might converge on HVC neurons that also receive cancellation signals from the inhibitory network, thus providing a discrete cellular site where the three signals could be combined. Alternately, auditory representations of the model could enter at later stages of the circuit, perhaps in the AFP. Although current thinking holds that HVC serves as the only source of auditory input to the AFP, other possible points of auditory entry include the thalamic nucleus DLM or area X. Therefore, future research should try to define the number of distinct auditory inputs to HVC and to determine whether auditory information independently enters song system structures other than HVC.

A final concern is the use of song playback to probe the location of a song template. An assumption that we and others have made is that the neuronal selectivity for song playback somehow reflects the memory of that song. In the case of the BOS this idea seems reasonable, but the BOS is also a song that the bird currently performs, rather than a song that exists only in memory. This motor “confound” has made it difficult to determine the extent to which BOS selectivity purely reflects auditory memory.49,50 Indeed, extracellular studies that have used tutor song playback to probe auditory responses in the song system have as yet failed to present convincing evidence of a tutor song bias.49−52 In the experiments we have undertaken, the BOS rather than the tutor song has served as the probe stimulus, and thus the synaptic activity evoked by the tutor song in HVC remains unknown. One expectation of a memory is that an activity pattern resembling that evoked by the stimulus that shaped the memory should arise spontaneously, rather than depending on presentation of the stimulus itself.53 In the songbird, future studies must determine whether tutor song memories are captured in the patterns of either auditory-evoked or spontaneous activity.54

REFERENCES