Synaptic Transformations Underlying Highly Selective Auditory Representations of Learned Birdsong

Melissa J. Coleman and R. Mooney
Department of Neurobiology, Duke University Medical Center, Durham, North Carolina 27710

Stimulus-specific neuronal responses are a striking characteristic of several sensory systems, although the synaptic mechanisms underlying their generation are not well understood. The songbird nucleus HVC (used here as a proper name) contains projection neurons (PNs) that fire temporally sparse bursts of action potentials to playback of the bird’s own song (BOS) but are essentially silent when presented with other acoustical stimuli. To understand how such remarkable stimulus specificity emerges, it is necessary to compare the auditory-evoked responsiveness of the afferents of HVC with synaptic responses in identified HVC neurons. We found that inactivating the interfacial nucleus of the nidopallium (NIf) could eliminate all auditory-evoked subthreshold activity in both HVC PN types, consistent with NIf serving as the major auditory afferent of HVC. Simultaneous multiunit extracellular recordings in NIf and intracellular recordings in HVC revealed that NIf population activity and HVC subthreshold responses were similar in their selectivity for BOS and that NIf spikes preceded depolarizations in all HVC cell types. These results indicate that information about the BOS as well as other auditory stimuli is transmitted synaptically from NIf to HVC. Unlike HVC PNs, however, HVC-projecting NIf neurons fire throughout playback of BOS as well as non-BOS stimuli. Therefore, temporally sparse BOS-evoked firing and enhanced BOS selectivity, manifested as an absence of suprathreshold responsiveness to non-BOS stimuli, emerge in HVC. The transformation to a sparse auditory representation parallels differences in NIf and HVC activity patterns seen during singing, which may point to a common mechanism for encoding sensory and motor representations of song.

Key words: auditory; tuning; songbird; in vivo intracellular; vocal learning; communication; zebra finch; neuronal interaction; temporal sparseness

Introduction

Neurons with remarkably specific stimulus-response properties characterize higher order sensory areas of the CNS. These include face- and motion-sensitive cells in primate visual cortex (Maunsell and Newsome, 1987; Rolls, 1992; Sugase et al., 1999), whisker–selective neurons in rodent barrel cortex (Welker, 1976), pulse–echo-sensitive neurons in bat auditory cortex (Suga et al., 1978; Kawasaki et al., 1988; Fitzpatrick et al., 1993), and space-specific neurons in owl auditory midbrain (Knudsen and Konishi, 1978; Wagner et al., 1987). Perhaps the most selective auditory neurons yet described are in the songbird telencephalic nucleus HVC (used here as a proper name), a sensorimotor nucleus essential to learned vocalizations (see Fig. 1A). HVC projection neurons (PNs) fire sparse bursts of action potentials in response to auditory playback of the bird’s own song (BOS) and are essentially unresponsive to other acoustic stimuli (Margoliash, 1983, 1986; Lewicki and Arthur, 1996; Mooney, 2000). This stimulus specificity distinguishes HVC PNs from neurons in primary auditory areas of the avian telencephalon (Lewicki and Arthur, 1996). Elucidation of how BOS selectivity in HVC arises can identify the origins of this stimulus specificity and inform models of auditory-guided vocal learning, because different HVC PN types innervate either song patterning or vocal plasticity pathways (Dutart et al., 1998; Kubota and Taniguchi, 1998; Mooney, 2000).

In several systems, stimulus–selective sensory responses result from the convergence of multiple afferents with different response properties (Jagadeesh et al., 1993; Ferster et al., 1996; Vidyasagar et al., 1996; Kawasaki and Guo, 1998; Pena and Konishi, 2001; Brecht and Sakmann, 2002; Perez-Olive et al., 2002; Higley and Contreras, 2003; Monier et al., 2003). A major limitation to our understanding of how BOS-evoked responses are generated is that the full complement of HVC auditory afferents remains uncharacterized (Kelley and Nottebohm, 1979; Fortune and Margoliash, 1995; Boco and Margoliash, 2001; Cardin and Schmidt, 2004). Extracellular studies suggested that the interfacial nucleus of the nidopallium (NIf) functions as a source of auditory input to HVC but could not address whether NIf interacts cooperatively with other HVC auditory afferents to facilitate auditory-evoked action-potential activity (see Fig. 1B). Extracellular studies suggested that the interfacial nucleus of the nidopallium (NIf) functions as a source of auditory input to HVC but could not address whether NIf interacts cooperatively with other HVC auditory afferents to facilitate auditory-evoked action-potential activity (see Fig. 1B). Extracellular studies suggested that the interfacial nucleus of the nidopallium (NIf) functions as a source of auditory input to HVC but could not address whether NIf interacts cooperatively with other HVC auditory afferents to facilitate auditory-evoked action-potential activity (see Fig. 1B)."
A further challenge to understanding how auditory representations of the BOS are generated and transformed in the songbird’s brain is that HVC contains several distinct cell types (see Fig. 1B), including interneurons and the two different PNs. To better understand how BOS representations may be transformed between NIf and HVC, we compared BOS-evoked multiunit NIf activity with simultaneous subthreshold activity in individual HVC PNs and interneurons. We found that suprathreshold BOS selectivity in NIf closely paralleled subthreshold selectivity in individual HVC neurons, indicating that information about the BOS and other stimuli is transmitted synthetically from NIf to HVC; however, individual HVC-projecting NIf neurons (NIfHVC) fired throughout BOS playback, unlike the phasic BOS-evoked firing of HVC PNs. NIf neurons also fired to other stimuli that evoked only subthreshold responses in HVC. These differences in NIf and HVC firing patterns suggest that temporally sparse BOS-evoked firing and an enhanced selectivity for the BOS arise in HVC. Because HVC PNs generate sparse activity during singing, these results may point to a common mechanism for encoding sensory and motor representations of song.

Parts of this paper have been published previously in abstract form (Coleman and Mooney, 2002).

Materials and Methods

General methods for these procedures have been described previously (Mooney, 2000; Rosen and Mooney, 2000).

Subjects. Experiments were performed using 88 adult (135–650 d after hatch) male zebra finches (Taeniopygia guttata) in accordance with a protocol approved by the Duke University Institutional Animal Care and Use Committee. All birds used in this study were bred and reared in our colony.

Stimuli. Before each experiment, songs were recorded from a male zebra finch placed in a sound-isolation chamber with a female zebra finch (Industrial Acoustics Company, Bronx, NY). Songs were amplified and low-pass filtered at 10 kHz, digitized at 20 or 22.05 kHz, and stored on a hard drive. Songs were recorded and edited using custom software (LabView, National Instruments, Austin, TX; written by M. Rosen, F. Livingston, and Ramani Balu). Edited songs included two or three motifs, the largest repeated unit in the bird’s song, and were typically 2–4 sec in duration. Stimuli that were presented included the following: (1) the BOS; (2) BOS played in reverse (REV), in which the temporal structure of individual syllables and the global syllable order are reversed; (3) BOS with the syllable order reversed (BOS-RO), in which the temporal order of individual syllables is preserved, but the syllable sequence is reversed; (4) song from a conspecific (i.e., zebra) finch (CON); and (5) white noise bursts, consisting of four to five bursts of white noise (each 50–100 msec in duration). Stimuli were presented at −70 dB sound pressure level (SPL), measured with a sound level meter (rms, A-weighted), with an interstimulus interval of 5 ± 1 sec.

Surgery. Animals were anesthetized with 20% urethane (60–100 μl total; Sigma, St. Louis, MO) administered into the pectoral muscle in 20–30 μl aliquots at 30 min intervals. In a few cases when needed, birds were administered halothane (Abbott Labs, Chicago, IL) via inhalation immediately before surgery. Lidocaine ointment (2.5%; Astra, Westboro, MA) was applied to the scalp, after which the scalp was dissected along the midline. The approximate x–y location of NIf was marked on the skull using stereotaxic coordinates (provided by Dr. J. M. Wild, University of Auckland, Auckland, New Zealand). Briefly, the marking pipette was set at interaural zero and then moved 0.3 mm caudal to the bifurcation of the midsagittal sinus by rotating the head angle of the bird in the stereotaxic apparatus. The skull overlying NIf was then marked at a position 1 mm rostral and 1.7 mm lateral from the bifurcation. In addition, the skull overlying HVC was marked at 0 mm rostral and 2.4 mm lateral to the bifurcation. A stainless steel post was then attached to the rostral part of the bird’s skull with dental cement and cyanoacrylate. Once the cement hardened, the bird was removed from the stereotaxic device and placed in a sound-attenuating chamber (Industrial Acoustics Company) on an air table (TMC, Peabody, MA). The bird’s head was immobilized via the mounted post, and its body temperature was maintained via an electric blanket set at 37°C (Harvard Apparatus, Holliston, MA). All recordings were made at a 50–55° head angle in the recording chamber. Small craniotomies were made over HVC and NIf, and the dura was slit open with an insect pin. Recording electrodes were lowered into NIf and HVC using a one-dimensional hydraulic micromanipulator (Soma Scientific, Irvine, CA). NIf typically was encountered at a depth of 1.8–2.2 mm and HVC at a depth of 0.3–0.8 mm from the surface of the brain. The NIf electrode was always lowered into the brain at a vertical orientation; to perform simultaneous recordings from the ipsilateral NIf and HVC, the HVC electrode was lowered into the brain at 15° below vertical.

Electrophysiology and song presentation. For extracellular recordings, multiunit activity in NIf was recorded using either glass micropipettes or carbon fiber electrodes (0.4–0.8 MΩ; Kation Scientific, Minneapolis, MN). Glass electrodes were pulled on a vertical puller to have a long taper, and the tip was broken to a diameter of 10–30 μm. The electrodes were tip-filled with 0.5% tetramethylrhodamine dextran (3000 MW; Molecular Probes, Eugene, OR) in 1 M NaCl to facilitate post hoc visualization of the electrode tip. The remainder of the electrode was filled with 1 M NaCl. Extracellular signals were amplified via an A-M Systems model 1700 differential amplifier (Everett, WA) and bandpass filtered between 300 Hz and 5 kHz. Intracellular recordings were made with sharp electrodes (100–180 MΩ; borosilicate glass, BF100–50; Sutter Instruments, Novato, CA) filled with 2 M KAc. All intracellular electrodes were first tip-filled with 5% neurobiotin in 2 M KAc. Recordings were amplified via an AxoClamp 2B amplifier (Axon Instruments, Foster City, CA) and low-pass filtered at 3 kHz. All recordings were digitized at 10 or 11.025 kHz and stored on a PC. All intracellular recordings were obtained in bridge mode, in which the bridge was balanced by offsetting the electrode series resistance to 200 pA, 50 msec current injections before the cell was impaled. The bridge was checked periodically and rebalanced as needed during the recording session.

NIf was identified based on its electrophysiological properties and its response to playback of BOS and REV (see Fig. 2). As noted previously (Janata and Margoliash, 1999) and described here, NIf produces spontaneous bursts of activity and responds more to presentation of BOS than REV, two features that distinguish it from the surrounding tissue. Initially, placement of the recording electrode in NIf was confirmed anatomically, by iontophoretic (+5 μA alternating on/off current, for 20 min) dye injections or by electrolytic lesions (−10 μA for 15 sec) at the recording site. Once NIf was located, intracellular recordings were obtained from individual HVC neurons that were identified based on their firing properties and response to current injection (Dutar et al., 1998; Kubota and Taniguchi, 1998; Mouney, 2000). Intracellular recordings from NIf neurons were made after first locating NIf with a carbon fiber electrode. In several (n = 5 of 14) of the NIf intracellular recordings, the location and identity of the neuron were confirmed by intracellular staining with neurobiotin and subsequent visualization of a cell body within the boundaries of NIf and axon terminals extending into HVC (i.e., NIfHVC neurons). We believe that the entire population of NIf neurons that we sampled here were most likely NIfHVC neurons because their spontaneous, DC-, and song-evoked firing patterns were similar to the identified NIfHVC neurons and because intracellular recordings made in vitro are heavily biased to projection neurons (our personal observations). In addition, the response strength (RS) of identified NIfHVC neurons and the RS of the unidentified NIf neurons were not statistically different (see Results and Fig. 12A).

NIf inactivation. To inactivate NIf, GABA (250 mM in 1 M NaCl; Research Biochemicals, Natick, MA) was pressure ejected with a Picospritzer (15–40 msec pulses at 10–20 psi; General Valve, Fairfield, NJ) into NIf through the NIf recording electrode. The location and extent of GABA application was inferred post hoc from the distribution of rhodamine dextran (3000 MW) that was added to the GABA solution for a 0.5% final concentration (see Fig. 3C). One experiment was excluded in which subsequent histological analysis revealed that there was no rhodamine in the recording area, suggesting that no GABA was ejected from
the pipette. The effect of GABA on NIF activity was monitored through the same electrode that contained the GABA solution.

In some NIF inactivation experiments, HVC action-potential activity was monitored with an extracellular carbon fiber electrode. In these experiments, at least 20 iterations of BOS were presented before GABA was ejected from the recording electrode. After a baseline response was established, GABA was infused into the preparation between each stimulus presentation for three to five iterations of the presented auditory stimuli (i.e., once every 5–10 sec). Auditory presentations continued without GABA application until evoked HVC activity recovered to pre-GABA levels. The response strength (see below) was calculated for each BOS presentation before, during, and after GABA application. Control injections consisted of 0.5% rhodamine dextran in 1 mM NaCl.

In other NIF inactivation experiments, GABA application to NIF was paired with intracellular recordings from HVC neurons. In these experiments, at least 10 iterations of each auditory stimulus were presented before GABA was applied to NIF. GABA was then applied to NIF between each auditory presentation for three to five iterations of the entire set of auditory stimuli. The effect of NIF inactivation on HVC responses was calculated off-line in MatLab (The MathWorks, Natick, MA) in which the integral of the subthreshold response to each BOS playback was calculated, relative to the baseline resting membrane potential. In all of these calculations, action potentials were removed from the subthreshold response by clipping the spikes near their threshold (the point at which the membrane potential became regenerative). To normalize the responses, the HVC neuron response for each BOS presentation during and after GABA application in NIF was divided by the mean pre-GABA responses (10–11 iterations) to obtain the percentage change in HVC neuron response. The maximum percentage change of each neuron to BOS presentation after GABA application was used to calculate the mean maximum change for all experiments.

Data analysis. Many aspects of the analysis of intracellular data use methods described previously (Mooney, 2000; Rosen and Mooney, 2003). NIF activity was quantified off-line in LabView. The threshold for detecting units within NIF multiunit activity was set visually by the user at an amplitude above background that excluded the smaller amplitude events (see Fig. 2). Our aim for setting a high threshold was to include events from larger neurons and cell bodies and to exclude the activity of fibers of passage. Although we did not analyze the number of neurons that contributed to the multiunit activity, we estimate that these multiunit recordings sample the activity of 5–10 individual NIF neurons.

The NIF multiunit activity and the action potential response of HVC neurons were calculated using the RS, which is the difference in mean firing rate during the stimulus and the mean firing rate during a prestimulus baseline period of similar duration. A similar approach was used to measure subthreshold responses from intracellular recordings in HVC, where the positive area measured between the median-filtered average membrane potential during the prestimulus period and the baseline mode was subtracted from the positive area measured between the stimulus-related waveform and the baseline mode. Similar measurements were made to estimate the magnitude of song-evoked hyperpolarizations in HVCx neurons (Mooney, 2000, his Fig. 9). The responses for either firing rate or subthreshold activity were expressed as z-scores. The z-score takes the difference between the response during the stimulus and the baseline response and divides that by the SD of that difference and is calculated by the following:

$$z = \frac{\bar{S} - \bar{B}}{\sqrt{\text{Var}(S) + \text{Var}(B) - 2\text{Covar}(S, B)}}$$

where $\bar{S}$ is the mean response during the stimulus, $\bar{B}$ is the mean response during a baseline period, and the denominator is the SD of $(S - B)$. For subthreshold responses, positive area z-scores refer to depolarizing responses, and negative area z-scores refer to hyperpolarizing responses.

One method to quantify neuronal selectivity for one stimulus versus another is the $d'$ value, which provides a statistical measure for the discriminability between two stimuli (Green and Swets, 1966). The $d'$ value is computed by the following equation:

$$d' = \frac{2(\bar{R}_{\text{STIM}} - \bar{R}_{\text{STIM2}})}{\sqrt{\sigma^2 \text{STIM1} + \sigma^2 \text{STIM2}}}$$

where $R$ is the response strength to the stimulus (STIM), $\bar{R}$ is the mean value of $R$, and $\sigma^2$ is its variance. For our analyses, the selectivity for BOS (STIM1) was compared with each of three stimuli: REV, BOS-RO, and CON (STIM2). A $d' > 0.5$ was used as the criterion for deeming a response selective (Solis and Doupe, 1997).

One limitation of using $d'$ values to assess relative responsiveness is that they are sensitive to response variability as well as response strength. Therefore, to compare the response of a neuron to two different stimuli without incorporating measures of variability, we also used the selectivity index (Solis and Doupe, 1997) as follows:

$$d' = \frac{\bar{R}_{\text{STIM}}}{(\bar{R}_{\text{STIM1}} + \bar{R}_{\text{STIM2}})}$$

Neurons preferring stimulus 1 (STIM1, i.e., BOS) over STIM2 (i.e., REV, CON, or BOS-RO) will have a similarity index near 1, those preferring STIM2 over STIM1 will have an index near 0, and those with no preference for either stimulus will have a value of 0.5.

Spike-triggered averages. Spike-triggered averages (STAs) were calculated to measure the relative timing of auditory-evoked changes in NIF multiunit activity and changes in HVC neuron membrane potential. STAs were calculated off-line in LabView from paired NIF and HVC neuron recordings during both song- and noise-evoked auditory responses and white noise bursts. Briefly, STAs were calculated by averaging the median-filtered HVC neuron membrane potential relative to NIF multiunit activity (NIF units) exceeding a user set trigger threshold (i.e., trigger events). STAs were calculated for a time window ±200 msec relative to the trigger event. STAs during song presentation were calculated from NIF activity elicited during playback of the entire stimulus. To correct for stimulus coordination artifacts, we shuffled NIF and HVC records from different trials and generated a shuffled STA that was then subtracted from the raw STA to generate a corrected STA that captured the trial-to-trial correlation at the two sites (Perkel et al., 1967; Ts’o et al., 1986).

A more accurate assessment of timing of auditory-evoked activity in NIF and HVC requires measuring stimulus-evoked response onset latencies. Such measurements are made more complicated by using song, a complex acoustical stimulus, during which the spike discharge is not reliably predicted by the stimulus onset. Therefore, we used white noise bursts to measure absolute evoked onset latencies in NIF and HVC and noise-evoked STAs to more accurately measure the relative timing of auditory-evoked activity in HVC neurons. The stimulus STAs of HVC neuron membrane potential were calculated from NIF multiunit activity that occurred within a narrow (~3–5 msec) time window 18–25 msec after the onset of white noise bursts. To determine the exact placement of the window, we first estimated the onset time for a given NIF multiunit recording by binning (bin size = 1 msec) all events that exceeded the user set trigger threshold. The largest bin (i.e., the most frequently occurring onset time) was then used to center the trigger window for calculating STAs. The onsets of STA peaks were estimated with a cumulative sum analysis. For this analysis we calculated the cumulative sum of the mean STA from 20 msec before and 30 msec after the NIF trigger event and designated the peak onset to be 5% of the maximum response. The cumulative sum did not reach a plateau, because the membrane potential did not return to baseline levels within the STA window. Therefore, we designated the maximum cumulative sum value to be the value at 30 msec after the NIF trigger event, which included the peak and the initial decay of the post-stripe STA.

Temporal sparseness index. To compare the relative temporal sparseness of action potential activity of NIF and HVC neurons, we calculated a “temporal sparseness index.” To calculate a sparseness index for each neuron, we first constructed peristimulus time histograms (PSTHs) (25 msec bins) and then determined the fraction of bins occupied with spikes during song presentation that exceeded the mean prestimulus activity level (average number of spikes per bin). This fraction was then subtracted from 1 to yield the index. Thus, a neuron that fired action potentials throughout song would have a sparseness index of 0, whereas a
neuron that fired only very sparsely would have a sparseness index approaching 1. In addition to calculating a temporal sparseness index for HVC neurons recorded during simultaneous NIf and HVC experiments, we also used data from previous intracellular recordings of identified HVC neurons from 20 animals, 19 of which were adults and 1 of which was a 72-day-old juvenile. Some of these data have been used in a previous publication (Mooney, 2000).

**Histology.** After each recording session, birds were deeply anesthetized with equithesin and transcardially perfused with 0.9% saline, followed by 4% paraformaldehyde (PFA) in 25 mM NaPO₄ buffer. Brains were removed from the skull and postfixed in 4% PFA with 30% sucrose overnight at 4°C, blocked sagittally, and sectioned on a freezing microtome at 50 μm. Neurobiotin-filled neurons were further processed by incubating the sections in avidin–Alexafluor 488 (Molecular Probes) and visualized using epifluorescence illumination. Extracellularly applied rhodamine dextran was also visualized using epifluorescence. To determine the position of the NIf recording site, drawings were made of the lesions from cresyl violet-stained sections or rhodamine labeling. Once these drawings were made, a composite schematic of all NIf recording sites was drawn, as shown in Figure 2D. To determine the location of rhodamine labeling relative to NIf location after GABA application, photomicrographs or digital images of the rhodamine labeling were superimposed on images of the same section viewed under combined dark-field and fluorescent illumination. In some cases, images of rhodamine labeling were superimposed on images of the same section after it had been stained further with cresyl violet to delineate the boundaries of NIf.

**Results**

A necessary step in understanding how selective auditory representations of the BOS are generated in HVC involves a comparison of activity in its auditory afferent(s) with subthreshold activity in single HVC neurons. To date, the degree to which a single auditory afferent can account for auditory-evoked subthreshold activity in HVC neurons remains unclear because direct comparisons of the activity in a candidate auditory afferent and subthreshold activity in the various HVC neuron types have not been made. We therefore (1) characterized auditory-evoked suprathreshold activity in NIf and compared across birds the BOS selectivity of NIf responses with subthreshold activity in identified HVC neurons, (2) tested the necessity of NIf activity for auditory-evoked suprathreshold and subthreshold activity in identified HVC neurons, (3) directly compared NIf multiunit activity with subthreshold activity in identified HVC neurons within birds, and (4) measured the BOS-evoked activity of identified NIf neurons that project to HVC.

**NIf auditory response properties**

Of the direct anatomic inputs to HVC, NIf is the most likely source of auditory input to HVC (Janata and Margoliash, 1999; Cardin and Schmidt, 2004) (Fig. 1B). To examine the degree to which the auditory responses of NIf can account for the subthreshold responses in HVC neurons, we first recorded NIf activity with a multiunit extracellular electrode and presented BOS, temporal variants of BOS (REV and BOS-RO), and conspecific song (CON). A representative recording (Fig. 2A) displays three general features of the NIf response: (1) excitatory responses to a wide range of auditory stimuli, including the songs of other zebra finches, (2) a stronger response to BOS than all other song types presented, and (3) a BOS-evoked response that was sustained over much of the stimulus duration. These responses are qualitatively similar to auditory-evoked subthreshold activity in HVC neurons, especially those that project to the song premotor nucleus RA (HVCRA) (Mooney, 2000).

The general features of NIf multiunit activity for this exemplar were representative of the NIf multiunit activity across birds. Playback of BOS and BOS-RO always elicited a significant excitation of NIf (80 of 80 NIf recordings for BOS and 74 of 74 NIf recordings for BOS-RO; *p* < 0.05) (see Fig. 2B). In addition, playback of CON elicited a significant excitatory response in 78% (57 of 73) of the NIf recordings, and REV elicited a significant response in 66% (53 of 80) of the NIf recordings. The mean BOS-evoked response was significantly greater than that evoked by any other song type (*p* < 0.05; ANOVA; *n* = 80 recordings in 41 animals) (Fig. 2B). The mean response strength to both BOS and BOS-RO (23.9 ± 1.8 and 18.7 ± 1.7 units/sec, respectively) was greater than that to either reverse REV OR CON (5.6 ± 0.7 and 8.1 ± 0.96 units/sec, respectively; *p* < 0.05; Tukey HSD), which did not differ significantly from each other (*p* > 0.05; Tukey HSD) (Fig. 2B). Although, like HVC, NIf responds more to BOS than to the other auditory stimuli, the high proportion of sites excited by CON and REV distinguishes NIf from HVC and other song-related brain nuclei (Margoliash, 1986; Doupe and Konishi, 1991; Margoliash et al., 1994; Lewicki and Arthur, 1996).

**Population comparisons of NIf activity and subthreshold activity in HVC**

The response strength measurements indicate that NIf neurons at the population level are selective for BOS over the other stimuli tested (i.e., NIf is BOS selective). We were curious to know the extent to which the suprathreshold BOS selectivity seen in NIf paralleled the selectivity of subthreshold responses seen in HVC. We first focused on HVCRA neurons because their subthreshold responses to acoustical stimulation are depolarizing from the resting membrane potential and therefore more straightforward to analyze than those in other HVC cell types (Mooney, 2000). The *d*’ value provides one method by which the selectivity for...
BOS over other stimuli can be quantified (Solis and Doupe, 1997) (see Materials and Methods). Here, we calculated the $d'$ value for BOS over REV for NIf multunit activity and compared those values with the subthreshold responses of HVC$_{RA}$ neurons measured with intracellular recordings [these HVC$_{RA}$ $d'$ values are comparable with those reported in Mooney (2000)]. Both NIf multunit activity and subthreshold responses in HVC$_{RA}$ neurons were very selective for BOS over REV, as indicated by $d' > 5$ (Fig. 2C) (NIf $d'_{NIf} = 2.56 \pm 0.14$, $n = 80$; subthreshold $d'_{HVC_{RA}} = 3.65 \pm 0.43$, $n = 20$). In addition, the $d'$ value for the HVC subthreshold response was significantly larger than the NIf $d'$ value for NIf multunit activity ($p < 0.05$; ANOVA; Tukey HSD). We also compared subthreshold $d'$ values of HVC$_{RA}$ neurons with $d'$ values of multunit activity in brain regions surrounding NIf. The regions surrounding NIf, on average, were significantly less selective for BOS over REV than regions in NIf or HVC (Fig. 2C) (NIf surround $d' = 0.27 \pm 0.27$; $n = 22$; $p < 0.05$; ANOVA; Tukey HSD). These indirect, across-bird comparisons suggest that the suprathreshold selectivity in NIf, rather than regions immediately surrounding NIf, more closely matches the subthreshold selectivity in HVC.

In initial experiments, we confirmed the anatomic location of our NIf recording sites by marking them with either electrolytic lesions or iontophoretic injections of rhodamine dextran ($n = 32$). At each recording site, we assessed the auditory selectivity for BOS over REV using the $d'$ metric and then superimposed the $d'$ score for a given site on a composite anatomical diagram of the regions from which we recorded (Fig. 2D). A notable feature of sites within the anatomically defined borders of NIf compared with adjacent brain regions was the selectivity for BOS over REV, reflected in the observation that nearly all recordings sites (29 of 32) within NIf had a $d' > 1$. These results suggest a sharp increase in auditory selectivity in NIf from surrounding auditory regions, including field L.

NIf is necessary for HVC auditory activity

To test directly whether NIf activity was necessary for auditory-evoked action-potential activity in HVC, we reversibly inactivated NIf with GABA application while monitoring HVC activity with an extracellular recording electrode ($n = 7$ NIf inactivations in five animals). As demonstrated recently (Cardin and Schmidt, 2004), we found that NIf inactivation was accompanied by a marked reduction or abolition of all auditory-evoked action-potential activity in HVC (Fig. 3A,B). Overall, the BOS-evoked HVC response strength decreased 93–99% (mean ± SD, 95.9 ± 2.5%) immediately (< 2 sec) after GABA application to NIf. The effects of NIf inactivation on HVC were reversible, and the auditory-evoked spiking in HVC reappeared within 5–60 sec after GABA application to NIf was discontinued. In addition to abolishing auditory-evoked activity in HVC, GABA application to NIf also abolished the spontaneous bursts seen in HVC during control recordings (data not shown). In contrast, control injections of vehicle only (saline and rhodamine dextran) into NIf did not significantly alter BOS-evoked response strength in HVC ($n = 3$ sites in two birds; $p = 0.09$; paired t test). These data indicate that NIf provides excitatory input to HVC and that NIf activity is necessary for auditory-evoked spiking in HVC.

To assess the approximate anatomical location of focus of GABA application, we mixed rhodamine dextran with the GABA solution. Rhodamine labeling was distributed within NIf and extended to its dorsal border (Fig. 3C,D, left topmost trace for experiment in Fig. 3A; Fig. 3D for all inactivation experiments). Although labeling sometimes extended beyond the dorsal border of NIf, because of the orientation of our recording–inactivation electrode, we never detected auditory-evoked activity in this region. These results suggest that the abolition of HVC auditory activity is caused by GABA application in NIf and not by spread to surrounding regions outside of NIf.

NIf is the main source of auditory input to both HVC projection neuron types

Although GABA inactivation of NIf reduces both spontaneous and auditory-evoked action potential generation in HVC, NIf may act permissively with other HVC auditory afferents to enable auditory-evoked spiking in HVC and generate song-selective responses. In addition, the extracellular HVC recordings that we made here did not allow us to distinguish between the different HVC neuron cell types, which could receive auditory input from
neurons that innervate a subregion of the basal ganglia homolog, area X (HVCX neurons), part of a pathway essential to vocal plasticity, show a mixture of hyperpolarizing and depolarizing responses to BOS playback, unlike the purely depolarizing responses evoked in HVCRA neurons (Mooney, 2000; Mooney et al., 2001). Thus, it was necessary to determine the effect of NIf inactivation on song-evoked subthreshold activity in identified HVC neuron subtypes. To examine the contribution of NIf activity to auditory-evoked activity in both HVC PN types, we measured synaptic activity in identified HVC neurons with intracellular electrodes and applied GABA to NIf (n = 5 HVCRA neurons in four animals; n = 4 HVCX neurons in four animals) (Fig. 4). In the most dramatic case, GABA application completely and reversibly abolished spontaneous activity in identified HVC neurons (Fig. 4). In the four other HVCRA Neurons, BOS-evoked synaptic activity also was greatly reduced, although not completely abolished (Fig. 4 B) (mean maximum decrease for all experiments = 70.7 ± 10.2%; n = 5). NIf inactivation also abolished or greatly reduced synaptic activity in HVCX neurons (Fig. 4 C, D). In three of these cells, hyperpolarizing current was injected into the cell to render the BOS-evoked response entirely depolarizing, simplifying comparisons of activity between the two areas. In one of these cells, GABA application in NIf resulted in an almost complete abolition of spontaneous and BOS-evoked activity in HVCX (Fig. 4 C); the response of this neuron recovered to 80.5% of the pre-GABA level after GABA application was discontinued. In the other two cells, BOS-evoked synaptic activity in HVCX was reduced but not abolished during GABA application to NIf (Fig. 4 D). In these three cases, there was a 64.7 ± 12.2% mean maximum decrease in synaptic response during NIf inactivation. In the fourth HVCX neuron, the membrane potential was maintained at rest (approximately −70 mV) and displayed a net hyperpolarizing response during BOS playback. In this case, GABA application in NIf abolished the BOS-evoked hyperpolarization. In sum, these experiments show that NIf activity is necessary for most or all of the spontaneous and song-evoked synaptic activity in both HVC PN subtypes and indicates that NIf is the major source of auditory input to HVC. In addition, these data suggest that both types of HVC projection neurons receive input (either directly or indirectly) from a common source, NIf.

**Direct NIf and HVC comparisons from simultaneous recordings**

To examine more closely the transformation of auditory activity between NIf and HVC, we compared directly the auditory activity in a small population of NIf neurons with subthreshold activity in HVC neurons by recording simultaneously from these areas. We initially focused on the relationship between NIf multunit activity and membrane potential changes in HVCRA neurons because their purely depolarizing patterns of subthreshold activity simplify comparisons (Mooney, 2000). We first give a brief qualitative description of these results and then provide a more quantitative assessment of the auditory-related activity in NIf and HVCRA neurons.

**Qualitative measures of auditory responsiveness in NIf and HVC**

Qualitatively, we noted that spontaneous bursts of NIf multunit activity were closely associated with bursts of PSPs in HVCRA neurons, although the HVCRA neuron usually showed little or no spontaneous action-potential discharge (Fig. 5 A). These spontaneous PSPs were preceded by NIf activity, as predicted from our inactivation experiments (Fig. 5 A, bottom right). As with spontaneous activity, NIf multunit activity paralleled the subthreshold activity in HVCRA neurons during auditory stimulation (Fig. 5 B). To examine the auditory-evoked activity in NIf and HVC, we presented BOS and temporally manipulated versions of BOS, as well as noise bursts, while recording simultaneously from NIf and HVCRA neurons (Figs. 5 B, 6). The correspondence between NIf activity and subthreshold activity in HVCRA is shown during playback of a single presentation of each auditory stimulus (Fig. 5 B) or multiple presentations of each stimulus (Fig. 6 A). These

---

**Figure 3.** NIf inactivation abolishes BOS-evoked action potential responses in HVC. **A**, Raw data showing the abolition of BOS-evoked activity in HVC after GABA application to NIf. **B**, Changes in HVC response strength during NIf inactivation. **C**, Percentages refer to the reduction in HVC BOS-evoked response during GABA application.
representative traces show a correspondence between auditory-evoked activity in NIF and subthreshold activity in HVC, but also reveal discordance in the types of auditory stimuli that could elicit action-potential activity in each area. Although only the BOS and BOS-RO routinely elicited action-potential responses in HVCRA neurons, all stimuli, including noise bursts, evoked action-potential activity in NIF (Fig. 2). The ability of white noise bursts to elicit action potentials in NIF and subthreshold activity in HVC neurons (Fig. 5B, rightmost recording) was useful in determining the auditory latencies in these two areas (see Figs. 9, 10).

Previous studies have shown that different subthreshold events underlie temporally sparse BOS-evoked firing patterns in the two HVC PN types (Mooney, 2000; Mooney et al., 2001), raising the possibility that the functional relationship between NIF and HVC varies with PN cell type. These different subthreshold responses involve sustained depolarization of HVCRA and net hyperpolarization of HVCX (Mooney, 2000). The subthreshold responses of HVCX arise from a combination of excitation and inhibition, the latter of which is thought to be driven by HVCInt (Mooney, 2000; Rosen and Mooney, 2003). Therefore, we also examined the relationship between NIF subthreshold activity and subthreshold activity in either HVCInt or HVCX (Fig. 6B,C). Like simultaneously recorded NIF multiunit activity, HVCInt was more strongly depolarized (Fig. 6A,C) in response to BOS and BOS-RO than to REV or CON. Furthermore, HVCX neurons had a larger hyperpolarizing response to BOS and BOS-RO than to REV or CON (Fig. 6B). These results are consistent with the idea that BOS-evoked hyperpolarizations in HVCX neurons are caused by feedforward inhibition from NIF through interposed HVCInt neurons.

Quantitative measures of auditory responsiveness in NIF and HVC

Quantitatively, NIF responsiveness to all song stimuli presented here paralleled the evoked subthreshold activity in HVC neurons. Using z-scores, we compared the responsiveness of simultaneously recorded NIF multiunit activity and subthreshold activity of identified HVC neurons with the playback of different song stimuli (Fig. 7A). When the z-score values for NIF multiunit activity were plotted against the z-score values of subthreshold responses for either HVCRA or HVCInt neurons, points clustered evenly around the unity line (Fig. 7A) (see Table 1 for statistical comparisons). Thus, within-bird comparisons reveal that the strength of auditory-evoked NIF multiunit activity parallels subthreshold activity for all auditory stimuli in HVCInt and for all stimuli except REV in HVCRA.

The relationship between song-evoked NIF activity and subthreshold activity in HVCX neurons depended on the HVCX neuron membrane potential. When HVCX neurons were near their resting membrane potential (−68.4 ± 1.5 mV), the BOS-evoked response was net hyperpolarizing (Fig. 6B), a feature most accurately calculated by the z-score value from the negative change in membrane potential area (negative area z-score) (Mooney, 2000). HVCX negative area z-score values at the resting membrane potential were significantly different from z-score values for simultaneously recorded NIF multiunit activity (Fig. 7A, bottom left, Table 1). This difference in z-score values could be attributable to additional inhibitory components from HVCInt to HVCX that mask excitatory inputs from NIF. To reduce the contribution of this inhibitory component, which is caused in part by a potassium current that reverses at approximately −85 mV (Rosen and Mooney, 2003), and enhance the excitatory component of the subthreshold response, we hyperpolarized the HVCX neurons (−84.3 ± 1.9 mV) and compared their positive area z-score values with z-score values of simultaneously recorded NIF multiunit activity. In contrast to the negative z-score values calculated near the resting membrane potential, the positive area z-scores calculated when HVCX neurons were hyperpolarized were not significantly different from z-scores of NIF multiunit activity (p > 0.05) (Fig. 7A, bottom right, Table 1). These results suggest that the difference in z-score values for NIF activity and subthreshold input in HVCX neurons near resting membrane potential could be attributable to contributions from synaptically interposed HVCInt neurons.

Within-bird comparisons of the selectivity for BOS over other stimuli in NIF and HVC

The comparisons of NIF multiunit activity and HVCRA subthreshold responses across different birds suggested an increase in selectivity between these two sites (Fig. 2C). This difference could be caused by variations in selectivity across birds, rather than a difference between sites within individual birds. To distinguish between these two possibilities, we simultaneously recorded NIF multiunit activity and HVC subthreshold activity and
calculated \( d' \) values at these two sites to compare the selectivity for BOS over other auditory stimuli (i.e., REV, BOS-RO, and CON). This comparison revealed a similar selectivity for BOS over REV and over CON in these two areas of the same bird (Fig. 7B). When the \( d' \) values for the NIf multiunit activity of BOS versus the three other auditory stimuli were plotted against the subthreshold \( d' \) value for HVC\(_{RA} \) and HVC\(_{Int} \) neurons, resultant points were distributed evenly about the unity line (mean \( d' \) BOS vs REV: HVC\(_{RA} \), 3.99 ± 0.43; NIF, 3.92 ± 0.52; paired t test; \( p = 0.904; n = 17 \); HVC\(_{Int} \), 2.89 ± 0.44; NIF, 2.89 ± 0.25; \( p = 0.996; n = 13 \); mean \( d' \) BOS versus CON: HVC\(_{RA} \), 3.65 ± 0.68; NIF, 3.07 ± 0.47; \( p = 0.316; n = 11 \); HVC\(_{Int} \), 2.61 ± 0.60; NIF, 2.24 ± 0.18; \( p = 0.552; n = 10 \); mean \( d' \) BOS versus BOS-RO: HVC\(_{RA} \), 0.96 ± 0.30; NIF, 1.13 ± 0.22; \( p = 0.516; n = 12 \); HVC\(_{Int} \), 0.93 ± 0.37; NIF, 0.72 ± 0.35; \( p = 0.573; n = 9 \). In contrast, when HVC\(_{X} \) neurons were at their resting potential, \( d' \) values calculated from NIf multiunit activity and subthreshold activity in HVC\(_{X} \) neurons were significantly different for all comparisons except BOS versus BOS-RO (mean \( d' \) BOS versus REV: HVC\(_{X} \), 2.26 ± 0.47; NIF, 3.03 ± 0.37, paired t test; \( p = 0.178; n = 20 \); mean \( d' \) BOS versus CON: HVC\(_{X} \), 2.21 ± 0.51; NIF, 2.84 ± 0.36; \( p = 0.226; n = 20 \); mean \( d' \) BOS versus BOS-RO: HVC\(_{X} \), 0.50 ± 0.26; NIF, 1.17 ± 0.20; \( p = 0.004; n = 21 \)). In this case, the \( d' \) value usually was larger in NIF than in HVC\(_{X} \), likely reflecting inhibition onto HVC\(_{X} \) from HVC\(_{Int} \) neurons (Mooney, 2000; Rosen and...
Mooney, 2003). Consistent with this idea, when HVCX neurons were tonically hyperpolarized, their subthreshold \( d' \) values were not significantly different from those calculated in NIf (mean \( d' \) BOS versus REV: HVCX, 2.88 ± 0.51; NIF, 3.52 ± 0.52; paired \( t \) test; \( p = 0.125; n = 15 \); mean \( d' \) BOS versus CON: HVCX, 2.42 ± 0.39; NIF, 3.23 ± 0.39; mean \( d' \) BOS versus BOS-RO: HVCX, 1.12 ± 0.30; NIF, 1.32 ± 0.30). Cumulatively, these data indicate that BOS selectivity in a small population of NIf neurons, quantified using \( d' \) values, equals the selectivity of the subthreshold activity in different HVC neuron types.

Correlations between NIf activity and membrane potential changes in HVC neurons

To examine correlations in NIf activity with changes in HVC membrane potential, we calculated the STA of the HVC neuron membrane potential relative to NIf trigger events (see Materials and Methods). As predicted from the average subthreshold responses (Fig. 6), STAs of HVCRA and HVCIInt neurons showed a net depolarizing membrane potential offset during playback of BOS and BOS-RO, whereas HVCX neurons showed a net hyperpolarizing offset (Fig. 8A, Table 2). In HVCRA neurons and HVCX neurons, the membrane potential offsets generated during playback of BOS and BOS-RO were significantly larger than the membrane potential shift elicited by playback of REV and CON (\( p < 0.05 \); ANOVA; Tukey HSD). In HVCIInt, the membrane potential offset to BOS was significantly greater than that to REV and CON (\( p < 0.05 \); ANOVA; Tukey HSD) but was not different from that to BOS-RO. The response to BOS-RO was not different from that of either REV or CON (\( p > 0.05 \); ANOVA; Tukey HSD).

In addition to a correlated shift in membrane potential during song playback, STAs also showed that NIf multiunit activity preceded membrane potential peaks in all HVC neuron subtypes. Song-evoked NIf multiunit activity preceded depolarizing peaks in HVCRA and HVCIInt neurons and mixed depolarizing and hyperpolarizing peaks in HVCX neurons (Fig. 8A). We subtracted any stimulus coordination artifact from the STA by first shuffling the BOS-evoked responses to generate a shuffled STA and then subtracting the shuffled STA from the raw song-evoked STA (Fig. 8B). On average, the STAs remaining after subtraction of the shuffled STA accounted for more than half of the STA peak depolarization elicited during BOS presentation in all cell types (HVCRA, 65.8 ± 7.6%, \( n = 11 \); depolarizing peak in HVCIInt with no current injection, 67.4 ± 6.3%, \( n = 14 \); hyperpolarizing trough in HVCX, 59.1 ± 11.1%, \( n = 7 \); HVCIInt, 71.0 ± 7.1%, \( n = 14 \)). These data indicate that NIf...
activity and subthreshold responses in HVC covary on a trial-by-trial basis.

Onset latencies of NIf and HVC auditory activity
To obtain an accurate assessment of the timing of auditory-evoked activity in NIf, we measured onset response latencies to short-duration (50–100 msec) white noise bursts (Fig. 9). Multiunit NIf activity was elicited at two distinct latencies after the noise burst onset: Gaussian fits of the responses were centered at 12.2 ± 0.27 and 19.3 ± 0.14 msec, respectively (Fig. 9B) (n = 52). Because multiunit recordings cannot distinguish between NIf neurons that project to HVC (NIfHVC) and other NIf subpopulations, we recorded intracellularly from NIf neurons and delivered noise bursts to the bird to define the auditory-evoked spike onset latencies of NIfHVC neurons. These neurons fired action potentials with a 20.4 ± 0.7 msec delay from the noise burst onset (range, 15–26; n = 14) (Fig. 9C), which overlapped with the longer latency peak of the noise-evoked NIf multiunit activity. Of the 14 NIf neurons from which we recorded, 5 were adequately filled to provide anatomical confirmation that they projected to HVC (i.e., NIfHVC neurons).

Previous cross-correlation analyses of extracellular activity in NIf and HVC have left unresolved whether NIf and HVC receive a common auditory input or form a serial auditory pathway (Janata and Margoliash, 1999). To more precisely estimate the relative timing of noise-evoked activity in these two areas, we paired extracellular recordings in NIf with intracellular recordings in HVC and calculated STAs of HVC neuron membrane potential from NIf trigger events. Given that the intracellular recordings that we made here showed that NIfHVC neurons had relatively long latency onset responses, we calculated the STAs from a narrow window around the longer latency component of the NIf multiunit activity (see Materials and Methods). In eight such simultaneous recordings of NIf multiunit activity and single HVCRA neurons, the resultant STAs were monophasic and positive, with a peak that followed the NIf trigger event by 6.5 ± 0.67 msec (Fig. 10). To determine the peak onset, we calculated the cumulative sum response of the mean STA during the time period shown in Figure 10 (−10 to 30 msec relative to NIf trigger event time), because this time frame captures both baseline and the peak response. We set the peak onset at 5% of the cumulative sum maximum (Fig. 10). Using this method, there was a 1.3 msec delay from the longer latency NIf trigger event to the onset of the HVCRA STA peak. Similar recordings from NIf and HVCX revealed a peak that followed the NIf trigger event by 5.8 ± 0.3 msec and a peak onset that followed the NIf trigger event by 1.9 msec (n = 7). Recordings from NIf and HVCX neurons near their resting potential resulted in STAs that were usually (five of six) biphasic, with an initial positive peak occurring after the NIf trigger event followed by a hyperpolarizing response (n = 6). The hyperpolarizing component is likely caused by inhibitory input from HVCint. To reduce the contribution of this inhibitory component, we calculated STAs of HVCX neurons held at more hyperpolarized membrane potentials (−84.3 ± 4.1 mV; n = 9). These STAs were monophasic, with a depolarizing peak after the NIf trigger event (7.3 ± 1.06 msec). The cumulative sum calculation indicated that the STA onsets occurred 1.6 msec after the NIf unit time, similar to the delay between NIf and HVCRA activity. In summary, recordings from all HVC cell types revealed that positive STA peaks followed NIf action potential activity, a feature that suggests that NIf provides either direct or indirect excitatory input to all HVC cell types in response to auditory stimulation.

Comparison of NIf multiunit and single-unit activity
Although individual HVC PNs fire temporally sparse BOS-evoked action potentials, multiunit NIf activity was sustained throughout much of BOS playback (Figs. 5B, 6A, B). The difference in output activity at these two sites could indicate that a temporally sparse representation of the BOS emerges in HVC; however, multiunit recordings from NIf could not resolve whether individual NIfHVC neurons fire action potentials throughout song playback or instead fire only at distinct times during BOS playback. In the latter case, the more sustained NIf activity seen with multiunit recordings may reflect the activity of several individual NIf neurons that fire at different times during BOS presentation. To determine the extent to which NIf multiunit activity reflected the action-potential activity of individual NIf neurons, we recorded intracellularly from NIf neurons (n = 14 neurons in seven animals). We found that individual NIf neurons did not fire at one particular time during BOS stimulation and that their activity qualitatively resembled NIf multiunit activity recorded from the same hemisphere (Fig. 11). A subset of these neurons (n = 5) was positively identified as NIfHVC neurons by neurobiotin staining and post hoc visualization of their labeled
Table 2. Membrane potential offsets in HVC neurons during song presentation

<table>
<thead>
<tr>
<th></th>
<th>BOS</th>
<th>n</th>
<th>REV</th>
<th>n</th>
<th>BOS-RO</th>
<th>n</th>
<th>CON</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVCRA</td>
<td>6.7 ± 0.55mV</td>
<td>9</td>
<td>2.3 ± 0.4 mV</td>
<td>9</td>
<td>5.2 ± 1.0 mV</td>
<td>9</td>
<td>2.6 ± 0.55mV</td>
<td>9</td>
</tr>
<tr>
<td>HVCx</td>
<td>-1.52 ± 0.55mV</td>
<td>13</td>
<td>0.38 ± 0.22mV</td>
<td>13</td>
<td>-1.49 ± 0.49mV</td>
<td>13</td>
<td>0.07 ± 0.26mV</td>
<td>12</td>
</tr>
<tr>
<td>HVCint</td>
<td>3.22 ± 0.56</td>
<td>14</td>
<td>0.85 ± 0.38</td>
<td>12</td>
<td>2.31 ± 0.49</td>
<td>14</td>
<td>1.28 ± 0.5</td>
<td>12</td>
</tr>
</tbody>
</table>

Individual NIf neuron output compared with individual HVC neuron output

Previous studies have suggested that selectivity for BOS over other auditory stimuli may be enhanced within HVC. Our comparisons of multunit NIf activity and subthreshold activity in HVC neurons suggest a close correspondence in selectivity (measured by the $d'$ metric) at these two sites, at least at the population level. One remaining issue, however, is whether the selectivity of NIf output at the single-unit level differs from the selectivity of individual HVC PNs and interneurons. We found that the mean $d'$ values for individual NIf neurons were not significantly different from the mean $d'$ values that we calculated for HVC PNs and interneurons ($d'$ BOS versus REV: NIf, 1.29 ± 0.23, $n = 14$; HVCRA, 1.53 ± 0.34, $n = 14$; HVCx, 0.59 ± 0.31, $n = 19$; HVCint, 2.54 ± 0.47, $n = 9$; $p > 0.05$; ANOVA; $d'$ BOS versus CON: NIf, 1.26 ± 0.23, $n = 12$; HVCRA, 1.72 ± 0.58, $n = 8$; HVCx, 0.68 ± 0.29, $n = 18$; HVCint, 0.27 ± 0.51; $p > 0.05$; ANOVA; $d'$ BOS versus BOS-RO: NIf, 0.25 ± 0.24, $n = 12$; HVCRA, 0.74 ± 0.34, $n = 9$; HVCx, 0.41 ± 0.21, $n = 19$; HVCint, 0.69 ± 1.23, $n = 8$; $p > 0.05$; ANOVA) (Fig. 12B). These comparisons indicate that BOS selectivity, as measured with the $d'$ value, does not differ between the output of NIf and the output of HVC.

The $d'$ metric incorporates both the difference in response strength to two stimuli and response variability (see Materials and Methods). Therefore, similar $d'$ values in NIf and HVC cannot be assumed to reflect similar response strength differences at the two sites. To determine the degree to which similar $d'$ values in NIf and HVC can be attributed to similar response strength differences, we also calculated a selectivity index for all cell types, in both NIf and HVC (see Materials and Methods) (Solis and Doupe, 1997). This index involves a within-cell comparison of responsiveness with two different stimuli without the potential confound of response variability. Similar to the $d'$ analysis, we found that there was no difference in the selectivity indices between NIf and HVC (BOS-REV: NIf, 0.75 ± 0.05; HVCRA, 0.86 ± 0.08; HVCx, 0.93 ± 0.16; HVCint, 1.27 ± 0.22; ANOVA; $p > 0.05$; BOS-CON: NIf, 0.54 ± 0.04; HVCRA, 0.75 ± 0.08; HVCx, 0.62 ± 0.12; HVCint, 0.69 ± 0.03; ANOVA; $p > 0.05$; BOS-BS-CON: NIf, 0.75 ± 0.04; HVCRA, 1.33 ± 0.32; HVCx, 0.81 ± 0.2; HVCint, 0.94 ± 0.08; ANOVA; $p > 0.05$). This result suggests to us that the similar $d'$ values measured in NIf and HVC are attributable to differences in relative response strengths rather than differences in response variability.

Although $d'$ values and selectivity index measurements do not differ significantly between the output of NIf and HVC, HVC PNs typically did not fire to non-BOS stimuli and thus appeared to have a different response to these stimuli, as measured by the $d'$ metric. This suggests that the differences in response strength measured in NIf and HVC are attributable to differences in response variability.

Axon terminals in HVC (data not shown). We compared the response strength of identified NIf-HVC neurons with that of the unidentified NIf neurons and found that these two groups were not significantly different for all auditory stimuli (unpaired t test; BOS, $p = 0.76$; REV, $p = 0.55$; BOS-RO, $p = 0.35$; CON, $p = 0.56$) (Fig. 12A). Because there was no significant difference in these two groups, we pooled data collected from all intracellularly recorded NIf neurons for the results presented in the following section.
strength of NIf neurons was significant for all stimuli tested (one sampled t test, compared with 0; \( p < 0.05 \)), whereas REV and CON often failed to evoke any action-potential responses in HVC neurons (one sampled t test, RS compared with 0; HVCRA: REV, \( p = 0.61 \); CON, \( p = 0.47 \); HVCX: REV, \( p = 0.08 \); CON, \( p < 0.05 \); HVCInt: REV, \( p = 0.12 \); CON, \( p = 0.17 \)). As predicted by comparisons of NIf multiunit activity and the action potential activity of HVC PNs, these results indicate that an enhanced selectivity for the BOS, characterized by a lack of responsiveness to non-BOS stimuli (i.e., REV and CON) emerges within HVC.

Another notable feature of the BOS-evoked response in HVC is its temporal sparseness, with some PNs firing only one action potential burst per song motif (i.e., the largest repeated song unit). To determine whether a temporally sparse response to the BOS also emerges within HVC, we calculated a temporal sparseness index for action-potential activity in both NIf and HVC (see Materials and Methods). We found that both HVC PN types had significantly greater sparseness indices than NIf neurons and that the temporal sparseness index for NIf and HVCInt were not significantly different (Fig. 12C) \( ( p < 0.0125 \); Wilcoxon test, corrected for multiple comparisons; 3 df). In sum, these data indicate that a temporally sparse response to the BOS emerges within HVC.

**Discussion**

We found that auditory-evoked NIf activity paralleled the subthreshold responses in individual HVC neurons and that NIf activity is necessary for the expression of subthreshold and suprathreshold auditory-evoked activity in HVC. Furthermore, NIf and HVC differ in their auditory-evoked firing patterns, with HVC PNs firing more sparsely to BOS playback and often not at all to non-BOS stimuli, whereas NIfHVC neurons fire throughout all auditory stimuli. These results indicate that an enhanced selectivity for the BOS emerges in HVC.

**NIf as the source of auditory input to HVC**

Historically, the identification of the auditory input of HVC has been difficult to resolve because of pronounced anatomical convergence from NIf, medial magnocellular nucleus of the anterior nidopallium (mMAN), nucleus uvaeformis (Uva), and possibly field L, the primary auditory telencephalon of the bird. Of these various afferents, Uva does not display auditory responses (Wild, 1994), and mMAN likely derives its auditory input indirectly.
from HVC (Vates et al., 1996), leaving NIf and field L as the two potential sources of ascending auditory input to HVC. Field L makes only a sparse direct input to HVC, and the response properties of field L do not match the synaptic responses seen in HVC PNs (Fig. 2) (Lewicki and Arthur, 1996). We propose that NIf provides the major source of auditory input to HVC and that the contributions of field L to the auditory responses in HVC are likely to be indirect, either through NIf or via the caudalateral mesopallium [caudalateral hyperstriatum ventrale in Vates et al. (1996)].

Selective responses to specific sensory cues can arise from convergence of inputs originating from either multiple or single areas of the brain (Suga et al., 1978; Kawasaki et al., 1988; Heiligenberg, 1989; Carr and Konishi, 1990; Rolls, 1992; Fitzpatrick et al., 1993; Vidyasagar et al., 1996; Pena and Konishi, 2001; Brecht and Sakmann, 2002; Pena and Konishi, 2002; Perez-Orive et al., 2002; Higley and Contreras, 2003). The generation of BOS selectivity in HVC neurons could be the result of afferents from multiple sources that converge onto individual HVC neurons. Although previous reports have shown that NIf inactivation abolished spontaneous and auditory-evoked spiking in HVC (Boco and Margoliash, 2001; Cardin and Schmidt, 2004), these studies did not exclude the possibility that other auditory afferents contributed to the response of HVC but was required NIf activity to permit BOS-evoked spiking in HVC neurons. We found, however, that NIf inactivation removes almost all of the subthreshold activity in HVC neurons, indicating that NIf is the major auditory afferent. Although our inactivation experiments do not fully exclude the possibility that field L contributes to the auditory-evoked response in HVC, this possibility seems unlikely, because field L neurons that project near or into HVC are distributed diffusely across a large area (Vates et al., 1996). In addition, our experiments do not exclude the possibility that NIf acts presynaptically within HVC to gate the activity of other HVC afferents or that additional auditory inputs dependent on HVC activity (i.e., mMAN) also contribute to HVC auditory responses in control recordings.

Convergent input to individual HVC neurons could arise from multiple neurons within NIf. Two findings from the present study consistent with this model are the concordance between NIf multiunit activity and HVC neuron subthreshold responses to auditory stimuli and the increased response strength of NIf multiunit activity compared with NIf single-unit activity. Anatomical convergence of NIf axons on HVC neurons could play an important role in the expression of temporally sparse and enhanced BOS selectivity, perhaps through synaptic integration or coincidence detection at the postsynaptic membrane.

Auditory transformations from NIf to HVC

Previous studies comparing the spiking patterns of HVC with its putative auditory afferents have suggested that BOS selectivity is enhanced within HVC (Janata and Margoliash, 1999). Lewicki and Arthur (1996) suggested that BOS selectivity is an emergent property of the HVC network, arising either directly or indirectly from nonselective inputs from field L. Janata and Margoliash (1999) found that neurons in NIf were selective for BOS but that selectivity increased from NIf to HVC. In contrast, using BOS values, we observed that BOS selectivity for a small population of NIf neurons closely matched the selectivity of the population of subthreshold inputs to individual HVC neurons (Mooney, 2000). This correspondence indicates that HVC neurons act as followers of a highly selective input from NIf. Furthermore, a somewhat surprising result is that the selectivity of individual NIf neurons was nearly identical to the suprathreshold selectivity of individual HVC PNs, as measured with both BOS and a selectivity index. Unlike the BOS analysis, the selectivity index does not take into account response variability, and thus the similar selectivity in NIf and HVC can be attributed to similar suprathreshold response biases at the two sites.

Although BOS and selectivity index measurements show that the output of NIf and HVC are equally selective for BOS over other auditory stimuli, calculations of action potential response strength in these two areas show that HVC is a site where enhanced selectivity to self-generated vocalizations emerges: unlike NIf PNs, HVC PNs do not fire action potentials to REV or CON. This more exclusive response also is characterized by temporally sparse firing in both HVC PN types, a feature that emerges within HVC. A similar increase in sparseness and selectivity has been shown between the odor-evoked activity of the antennal lobe and mushroom body neurons in the locust (Perez-Orive et al., 2002; Theunissen, 2003). Our results indicate that significant changes occur in the auditory representations between NIf and HVC.

Studies in other sensory systems have shown that intrinsic membrane properties and local circuitry contribute significantly...
to the generation of selective responses (Vidyasagar et al., 1996; Bringuier et al., 1999; Zhu and Connors, 1999; Ferster and Miller, 2000; Vinje and Gallant, 2000). Similar mechanisms may underlie the highly selective, temporally sparse BOS-evoked firing patterns of HVC PNs. BOS-specific firing in HVC PNs may result in part from simple thresholding mechanisms (i.e., an “iceberg” effect) or other more complex processes involving intrinsic membrane properties and inhibitory inputs from the HVC network. Thresholding of synaptic inputs in both the oval auditory system and visual cortex have been shown to play an important role in the generation of sensory selectivity (Cardanini and Ferster, 2000; Pena and Konishi, 2001, 2002). Consistent with this mechanism operating in HVC, the resting membrane potential of HVC \(_{RA}\) neurons is relatively hyperpolarized (approximately \(-80\) mV), and tonic depolarization of HVC \(_{RA}\) neurons can enable action-potential responses to REV and non-BOS stimuli (Mooney, 2000; our unpublished observations). Inhibition also has been implicated in the generation of stimulus specificity in other systems (Sillito, 1975; Fujita and Konishi, 1991; Ferster and Miller, 2000; Perez-Orive et al., 2002). Indeed, postsynaptic inhibition appears to play an important role in establishing a more exclusive action-potential bias in HVC \(_X\) neurons because disrupting G-protein-mediated inhibition in these cells increases their spiking response strength to nonpreferred stimuli, i.e., BOS-RO (Rosen and Mooney, 2003). The inhibition in HVC \(_X\) neurons most likely arises from HVCAIn neurons (Mooney, 2000; Rosen and Mooney, 2003). These interneurons could be excited directly by NIF or indirectly via HVC \(_{RA}\) neurons. A direct NIF effect is suggested by the close correspondence between non-BOS responses in NIF and interneurons and the similarly short onset latencies between NIF suprathreshold activity and subthreshold responses in all HVC cell types. In either case, BOS-exclusive responses in the two output pathways of HVC appear to be generated by transforming a common excitatory input to the two HVC PNs via two different local mechanisms: one is more dependent on intrinsic membrane properties and the other is more dependent on synaptic inhibition.

Functional relevance

The auditory-evoked responses in NIF and HVC \(_{RA}\) neurons may shed light on the synaptic processes underlying song production. Indeed, recordings from individual premotor (HVC \(_{RA}\)) neurons during singing show that these neurons fire at precise times during song (Hahnloser et al., 2003), a firing pattern similar to the response that these neurons have to playback of the BOS (Mooney, 2000). Although the function of such sparse firing in the songbird remains unknown, sparse firing patterns in other systems are thought to facilitate learning and memory (Golomb et al., 1998; Olshausen and Field, 1996; Hyvarinen and Hoyer, 2001). In addition, NIF is involved in song production (McCasland, 1987), and the activity of NIF \(_{HVCA}\) neurons during BOS playback resembles their activity in singing birds (Hahnloser and Fee, 2003). That is, like their response to auditory playback of BOS, individual NIF \(_{HVCA}\) neurons respond throughout a motif, whereas HVC \(_{RA}\) neurons fire a single burst of action potentials during a single song motif. These parallels give credence to the notion that a common mechanism gives rise to similar auditory and motor song representations in HVC.

References


Wild J (1994) Visual and somatosensory inputs to the avian song system via nucleus uvaformis (Uva) and a comparison with the projections of a similar thalamic nucleus in a nonsongbird, Columba livia. J Comp Neurol 349:512–535.