Singing-Related Neural Activity Distinguishes Four Classes of Putative Striatal Neurons in the Songbird Basal Ganglia

Jesse H. Goldberg and Michale S. Fee
McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts

Submitted 25 November 2009; accepted in final form 26 January 2010

Goldberg JH, Fee MS. Singing-related neural activity distinguishes four classes of putative striatal neurons in the songbird basal ganglia. J Neurophysiol 103: 2002–2014, 2010. First published January 27, 2010; doi:10.1152/jn.01038.2009. The striatum—the primary input nucleus of the basal ganglia—plays a major role in motor control and learning. Four main classes of striatal neuron are thought to be essential for normal striatal function: medium spiny neurons, fast-spiking interneurons, cholinergic tonically active neurons, and low-threshold spiking interneurons. However, the nature of the interaction of these neurons during behavior is poorly understood. The songbird area X is a specialized striato-pallidal basal ganglia nucleus that contains two pallidal cell types as well as the same four cell types found in the mammalian striatum. We recorded 185 single units in Area X of singing juvenile birds and, based on singing-related firing patterns and spike waveforms, find six distinct cell classes—two classes of putative pallidal neuron that exhibited a high spontaneous firing rate (>60 Hz), and four cell classes that exhibited low spontaneous firing rates characteristic of striatal neurons. In this study, we examine in detail the four putative striatal cell classes. Type-1 neurons were the most frequently encountered and exhibited sparse temporally precise singing-related activity. Type-2 neurons were distinguished by their narrow spike waveforms and exhibited brief, high-frequency bursts during singing. Type-3 neurons were tonically active and did not burst, whereas type-4 neurons were inactive outside of singing and during singing generated long high-frequency bursts that could reach firing rates over 1 kHz. Based on comparison to the mammalian literature, we suggest that these four putative striatal cell classes correspond, respectively, to the medium spiny neurons, fast-spiking interneurons, tonically active neurons, and low-threshold spiking interneurons that are known to reside in area X.

INTRODUCTION

The striatum is the main input nucleus of the basal ganglia (BG) circuit and is widely implicated in motor control and learning (Graybiel 2008). Inactivations or lesions of the striatum result in impaired motor learning while neuropsychiatric disorders such as obsessive compulsive disorder and Huntington’s diseases are associated with striatal dysfunction (Packard and Knowlton 2002). For decades, striatal function has been interpreted mainly in terms of the relatively large population of GABAergic medium spiny neurons (MSNs) that constitute striatal outputs (Hikosaka and Wurtz 1989). However, several interneuronal subtypes within the striatum are increasingly considered key components of BG function. Specifically, anatomical, histological, and physiological studies in brain slices indicate that the striatum contains at least four cell classes—MSNs, as well as three classes of striatal interneuron: fast-spiking interneurons, tonically active cholinergic neurons, and low-threshold spiking interneurons (Kreitzer 2009).

Surprisingly, the neural activity of all of these cell classes has not been described during a single behavior.

MSNs, the striatal output neurons, are the most numerous cell class (Wilson and Groves 1981). In brain slices, they are distinguished by their densely spiny dendritic arbor and their expression of the neuropeptides substance P or enkephalin (Mink 1996). During behavior, MSNs are distinguished by their low firing rates and their sparse neural activity that is often precisely time-locked to specific events and movements in a behavioral sequence (Barnes et al. 2005; Hikosaka and Wurtz 1989; Jin et al. 2009).

Fast-spiking interneurons (FS) are distinguished in brain slices by their high-frequency firing and their expression of the calcium-binding protein parvalbumin and of KV3-type potassium channels, which contribute to their uniquely narrow spike waveforms (Kubota et al. 1993; Lenz et al. 1994; Plenz and Kitai 1998). In vivo juxtacellular recording and labeling of FS neurons has confirmed that they are distinguishable by their narrow spike waveforms (Mallet et al. 2005), and this criterion has recently been used in the first study to characterize their activity during behavior. In contrast to MSNs, FS neurons exhibited constant firing (~20 Hz) and weak correlation to events in a well-learned maze task (Berke 2008).

Tonically active neurons (TANs), the cholinergic neurons of the striatum, are distinguished morphologically by their large cell bodies and expansive aspiny dendritic arbors (Zhou et al. 2002). Intracellular recording and labeling of TANs in vivo has demonstrated that they are distinguishable physiologically by their tonic activity and their long afterhyperpolarizations (Wilson et al. 1990). During behavior, TANs, in contrast to FS and MSN neurons (Berke 2008), do not generate high-frequency bursts and instead fire tonically at rates that can be modulated by a variety of events within a behavioral task, including reward, movement, and sensory stimuli (Aosaki et al. 1995; Apicella 2007; Morris et al. 2004).

Finally, perhaps the least understood of the striatal cell classes is the low-threshold spiking (LTS) interneuron. In brain slice studies, LTS neurons exhibit rebound low-threshold calcium bursts after release from hyperpolarization (Kubota and Kawaguchi 2000). To our knowledge, there are no reported recordings of LTS neurons in the behaving animal, although under anesthesia, putative LTS neurons reach the highest peak rate of all striatal neurons and discharge primarily in long bursts that contain multiple spikes (Sharott et al. 2009).

The goal of the present study was to record from all four of these striatal cell types during a BG-dependent behavior. To this end, the songbird presents an ideal model system. Songbirds have a specialized BG thalamocortical circuit, called the anterior fore-
brain pathway (AFP; Fig. 1A), that is required for song learning (Nottebohm et al. 1976; Scharff and Nottebohm 1991). Area X, the input nucleus of the AFP, contains the same four striatal cell classes that are found in the mammalian striatum (Fig. 1B) (Farries and Perkel 2002) as well as two pallidal cell classes that, as in mammals, can be distinguished from striatal cell types on the basis of their high spontaneous firing rate (DeLong 1972; Farries and Perkel 2002; Farries et al. 2005). The strong homology between the songbird and mammalian BG (Fig. 1B) is consistent with mounting evidence that the BG microcircuit has been conserved for >300 million years and predates the division of birds, mammals, and reptiles into separate classes (Reiner 2009).

We have recorded single neurons in area X of singing juvenile birds. Based on spike waveforms and singing-related neural activity, we find six distinct cell classes. In this study, we focus on the four cell classes that exhibited the low spontaneous firing rates (<30 Hz) that distinguish striatal from pallidal cell types (DeLong 1972; Johnstone and Rolls 1990). Notably, we find that the four putative striatal cell classes in the songbird area X exhibit singing-related firing patterns similar to the neural activities exhibited by their possible mammalian homologues: MSNs, FS interneurons, TANs and LTS interneurons. In addition, just as mammalian striatal neurons are correlated with the timing of ongoing behavioral tasks, all area X neurons exhibited significant correlations with the timing of specific syllables.

Our results provide a framework for the classification of extracellularly recorded striatal neurons in the songbird and include the first characterization of the singing-related neural activities of multiple cell types in a BG nucleus required for song learning.

**METHODS**

**Animals**

Subjects were 21 juvenile male zebra finches, 40–70 days post hatch (dph). Birds were obtained from the MIT zebra finch breeding
facility (Cambridge, MA). The care and experimental manipulation of the animals were carried out in accordance with guidelines of the National Institutes of Health and were reviewed and approved by the MIT Committee on Animal Care. All data are from juvenile birds singing undirected song.

**Chronic neural recordings and histology**

Recordings were carried out using a motorized microdrive described previously (Fee and Leonardo 2001). Units accepted for analysis had signal-to-noise ratios (average spike peak amplitude compared with SD of noise) >10:1. Data were acquired and analyzed using custom Matlab software (A. Andelman, D. Aronov and J. H. Goldberg). Small electrolytic lesions (20 μA for 15 s) were made through the electrodes at the conclusion of experiments for histological verification of electrode position.

**Data analysis**

Spikes were sorted off-line using custom Matlab software. We represented neural activities as instantaneous firing rates, \( R(t) \), defined at each time point as the inverse of the enclosed interspike interval (ISI) as follows (Eq. 1)

\[
R(t) = \frac{1}{t_{r+1} - t_r}, \quad \text{for} \quad t_r < t \leq t_{r+1}
\]

where \( t_r \) is the time of the \( r \)th spike. Peak firing rates (99th percentile rate) were calculated for each neuron as the inverse of its first percentile ISI. All ISI distributions were computed in bins that were evenly spaced in log time. Interbout periods were defined as silent periods separated from singing by \(<5\) s. Note that the interbout periods include the silent periods immediately preceding bout onset as well as following bout offset. Nonsinging firing rates were calculated from spiking activity during silent periods separated by singing by \(>10\) s. We used song to automatically trigger data acquisition as well as following bout offset. Nonsinging periods. We did not acquire data continuously during nonsinging periods. Thus the anecdotal observation that bursting in type-4 neurons may precede the onset of singing by tens of seconds is not supported by a quantitative analysis but is based on the recollection of the experimenters listening to neuronal activity on the audio monitor (Grass AM10).

**Burst definition**

While the classification of the four cell types did not depend on any definition of burst, we included a burst analysis to highlight differences in firing patterns that were not explicitly part of their classification scheme. Bursts were defined as a simple threshold-crossing of 250 Hz in the instantaneous firing rate (IFR). This threshold was based on the ISI distributions of the four cell classes—specifically that neurons types 1, 2, and 4 had peaks in their ISI distributions \(<4\) ms, whereas type-3 neurons did not (Fig. 7A). The observation that type-3 neurons did not burst was consistent with their relatively flat spike train autocorrelation functions (Fig. 7C).

**Spike width identification of type-2 neurons**

For each neuron, the spike width was calculated as the half-width of the negative going deflection of the average of 50 spike waveform examples, resulting in a total of \( 115 \) spike widths (115 neurons). The spike widths of type-2 neurons lay in a distinct cluster of 11 neurons separated by a gap 0.021 ms to the left of the edge of the remaining distribution. To determine the likelihood that such a distinct cluster could emerge from an underlying Gaussian distribution, we performed a Monte Carlo analysis where we generated 10,000 surrogate datasets of 115 spike widths, each drawn from a single Gaussian distribution with the same mean and SD as the observed dataset. From the surrogate datasets, we determined the probability of finding \( \geq 11 \) samples to the left of a gap of \(\geq 0.02\) ms. We found that in 0/10,000 datasets did such a condition exist. In fact, we found that even the observation of five samples to the left of a gap of 0.01 ms was extremely rare (\( P = 0.02 \)).

**Analysis of correlations of neural activity to song temporal structure**

For plastic song birds, syllables were identified both manually and with an automated clustering algorithm (A. Andelman) that uses standard acoustic features of zebra finch song (Tchernichovski et al. 2000). The songs of the plastic song birds in this study were not yet organized into motifs—thus we analyzed correlations on a syllable-by-syllable basis. We included in this analysis only those birds with song repertoires that included distinct, identifiable syllables and only analyzed cells that were recorded for \(>30\) syllable renditions (mean number of syllables per neuron = 169, range: 31–2,053). Neural activity from 60 ms prior to syllable onset to 60 ms following the median syllable duration was linearly time-warped, using as a reference the syllable’s median duration as described previously (Kao et al. 2008; Ölveczky et al. 2005). Syllables of subsong birds (\(<50\) dph) are highly unstructured, and neural correlations to their acoustic or temporal structure are outside the scope of this study.

To determine the significance of firing rate peaks and minima within syllables, a rate histogram (10-ms bin size) was generated of the syllable-aligned spike trains. Surrogate histograms were also calculated for all spike trains after a random time shift as described previously (Ölveczky et al. 2005). The shift was circular, such that spikes wrapped around to the beginning of the spike train. For each cell, the correlation distribution of the time-shifted firing rates was calculated with 1,000 different ensembles of random shifts. Peaks and minima in firing rate were considered significant when they crossed the 1st (for minima) and 99th (for peaks) percentile of the compiled minima and maxima of the surrogate histograms.

**Quantification of sparseness**

A “sparse” firing distribution of a neuron is one in which it gives a weak response during most stimuli (or movements), but a strong response during one stimulus (or movement) (Field 1994; Hahnloser et al. 2002). We used an “entropy” measure of sparseness (Lehky et al. 2005; Tolhurst et al. 2009) to quantify how selective neuronal activity was for specific times in the song. For each neuron, we calculated syllable-aligned rate histograms for every syllable (10-ms bin size as described in the preceding text), resulting in a total of \( N \) bins. These histograms were normalized over all syllables to generate a time-varying probability spike density function, \( p_i \), where the \( i \)th value indicated the normalized firing probability of that neuron within that time bin. Thus the sum of all points in this vector of length \( N \) was equal to 1 (\( \sum p_i = 1 \)). We then computed an entropy based sparseness index (Lehky et al. 2005) as follows (Eq. 2)

\[
\text{Sparseness}_\text{Index} = 1 + \frac{\sum p_i \log(p_i)}{\log(N)}
\]

The sparseness index is equal to 1 (maximal sparseness) when activity is restricted to a single portion of the song (a single 10-ms bin) and goes to zero when spikes are evenly distributed throughout the song.

**Trial-to-trial correlation analysis**

To quantify trial-to-trial differences in neural activity on short time scales, we computed the correlation coefficient (CC) between the instantaneous firing rate for all sets of trials, as follows (Eqs. 3 and 4),
RESULTS

Nonsinging firing rate distinguishes putative pallidal and striatal neurons in area X

We recorded 185 neurons in area X of 10 juvenile birds, aged 40–70 dph (see METHODS). In awake, nonsinging birds, we found that the average firing rate of area X neurons fell into two main classes: high-frequency (HF) neurons with average nonsinging firing rates >60 Hz (127 ± 30 Hz, n = 70, all error bars indicate SD unless otherwise stated), and low-frequency neurons with average nonsinging rates <30 Hz (4.9 ± 6.3 Hz, n = 115; Fig. 1D). We hypothesize that the high and low firing cells represent the pallidal and striatal neuronal cell classes, respectively, that are known to reside in area X (Fairies and Perkel 2002). This hypothesis is based on previous studies of area X neurons in brain slices, which revealed high spontaneous firing rates in pallidal neurons and low rates in striatal neurons (Fairies and Perkel 2002), and in vivo, which found that HF neurons in area X project to DLM (Leblois et al. 2009), as well as on numerous reports in mammals demonstrating high pallidal and low striatal firing rates in awake animals (Bar-Gad et al. 2003; DeLong 1972; Johnstone and Rolls 1990). In the present study, we focus exclusively on the putative striatal neurons.

Singing-related neural activity distinguishes four classes of putative striatal neuron in area X

Three measures of the neural activity divided the putative striatal neurons into four distinct classes: spike width, mean firing rate during singing, and peak firing rate (99th percentile rate, see METHODS) during singing. We first plotted spike width versus mean firing rate during singing and observed three distinct clusters. Type-1 neurons were defined by a firing rate <10 Hz and a spike width >0.06 ms (Fig. 1E, black cluster, n = 56). Type-2 neurons formed a distinct cluster of narrow spike waveforms (<0.06 ms; Fig. 1E, red cluster, n = 11, P < 0.0001, see METHODS). A third cluster had high mean firing rates during singing (>10 Hz) and broad spike waveforms (>0.06 ms; Fig. 1E, blue and green data points). Neurons within this cluster were bisected by their peak firing rates during singing (99th percentile rate, see METHODS)—type-3 neurons (Fig. 1F, blue cluster, n = 29) did not generate high-frequency bursts, exhibiting peak rates <600 Hz, while type-4 neurons (green cluster, n = 19) generated bursts at rates >600 Hz and frequently discharged at peak rates exceeding 1 kHz. We now consider the properties of each of these cell classes in turn.

Type-1 neurons exhibit sparse, temporally precise discharge during singing

A defining feature of type-1 neurons was their low average firing rate (Fig. 2). During nonsinging periods, type-1 neurons did not spike (<0.01 Hz, n = 56/56 neurons). During singing, spikes occurred more frequently (average rate = 1.58 ± 1.46 Hz), and occasionally discharged in burst bursts (Fig. 2, A and B). We defined bursts as events exceeding 250 Hz (see METHODS) and found that all type-1 neurons generated bursts that could reach high rates (peak firing rate within bursts = 480 ± 187 Hz; burst/s = 0.15 ± 0.12 s⁻¹, n = 56 neurons). The mean and peak firing rates exhibited by type-1 neurons during singing or nonsinging did not differ between subsong and plastic song birds (Table 1, P > 0.05, unpaired t-test, n = 22 and n = 34 neurons from subsong and plastic song birds, respectively).

In plastic song birds, where specific syllables could be identified (see METHODS), type-1 neurons exhibited sparse activity that was precisely locked to the occurrence of particular syllables or even specific times within single syllables (Fig. 2, A and B). We analyzed how neural activity was correlated with song timing in three ways. First, to quantify how spiking was distributed across the song, for each neuron, we computed an entropy-based sparseness index where a value of 1 indicated a neuron whose activity was entirely restricted to one portion of the song and a value of zero indicated a neuron that spiked evenly across the song (see METHODS). The activity of type-1 neurons was significantly more sparse than all other area X neurons (Fig. 2C, Table 1, sparseness index: type-1 = 0.45 ± 0.14, n = 31; types-2 to -4 = 0.023 ± 0.018, n = 25, P < 0.0001 unpaired t-test).

Second, to quantify the trial-to-trial variability in the firing rate modulations across multiple syllable renditions, we computed pairwise correlation coefficients (CCs) between smoothed spike trains of different song renditions (see METHODS). In all type-1 neurons recorded in plastic song birds, the timing of spikes was reproducible across repeated syllable renditions (CC = 0.26 ± 0.22, range: 0.02–0.85, n = 44 syllables from 31 neurons). We note, however, that even when neuronal activity was precisely time-locked to a given syllable, the number of spikes discharged was variable (Fig. 2E). On average, neurons spiking during less than half of the renditions of the syllable during which they spiked most reliably (39 ± 25%, range: 4.7–100%, Fig. 2E).

Third, to determine how neural activity was distributed within individual syllables, for each neuron we computed syllable-aligned rate histograms (as in Fig. 2B, bottom) and examined the temporal location of significant rate peaks (see METHODS). As a group, different type-1 neurons tended to spike at different times in the song (Fig. 2D) and at different times within syllables. Significant rate peaks were uniformly distribu-
FIG. 2. Type-1 neurons exhibit sparse, temporally precise singing-related activity. A: the raw voltage trace of a type-1 neuron and its instantaneous firing rate (see METHODS) are plotted beneath the spectrogram (age 64 dph). Note that this neuron spikes only during syllable “a” of a 3-syllable motif. B: top: expanded view of the voltage and spectrogram from the 1st motif from B (indicated by red bar). Middle: spike raster indicating spiking activity of this neuron during 73 renditions of the motif. Bottom: rate histogram compiled from the raster plot. C: top: for each putative striatal neuron recorded in a plastic song bird, the mean firing rate during singing is plotted against the sparseness index; color code as in Fig. 1 (see METHODS). The neuron from A and B is labeled with a solid black dot and an arrow. Bottom: histogram of sparseness index. D: spectrogram and spike raster of 6 type-1 neurons recorded in one bird (different bird from A and B) during ages 61–65 days post hatch (dph). Each neuron exhibits sparse activity temporally localized to distinct parts of a 3-syllable motif. E: cumulative probability plots of the number of spikes generated per syllable for all type-1 neurons recorded in plastic song birds. Red trace represents the average of 44 syllables from 31 neurons, and the dashed trace is from the neuron in A and B. Syllables that never contained a spike were excluded from this analysis. F: population histogram showing the distribution of significant rate peaks as a function of syllable timing. Data were compiled from 44 syllables from 31 type-1 neurons. Note that significant peaks occur throughout the syllables.

Thin-spiking, type-2 neurons exhibit brief high-frequency bursts throughout the song

Type-2 neurons were distinguished by their narrow spike waveforms (Fig. 1E). In contrast to type-1 neurons, type-2 neurons exhibited steady, low-frequency firing during nonsinging (7.9 ± 4.4 Hz, n = 11/11 neurons, Figs. 1D and 4A) with modest coefficient of variability in the duration of ISIs (CV_{ISI} = 1.36 ± 0.60; Fig. 4, B and C, Table 1). During singing, type-2 neurons exhibited higher firing rates and larger rate fluctuations (singing: mean rate = 19.1 ± 6.0 Hz, CV_{ISI} = 1.80 ± 0.37; P < 0.001 compared with nonsinging), including brief, high-frequency bursts (peak rate = 654 ± 51 Hz; burst/s = 2.27 ± 1.21 s^{-1}; Figs. 1, D and E, and 4, A–C). Bursting also occurred in the seconds immediately before and after song bouts (interbout silence, see METHODS) and was associated with a shift in the ISI distributions toward short intervals during singing and with a peak in the spike train autocorrelations at short times (Fig. 4, B and C). The firing patterns of type-2 neurons were indistinguishable in subsong and plastic song aged birds (Table 1, Fig. 4, n = 5 and 6 neurons from subsong and plastic song birds, respectively).

While type-1 neurons discharged sparsely at specific portions of the song, type-2 neurons discharged throughout the song (sparseness index = 0.044 ± 0.024) and exhibited sig-
Table 1. Spiking statistics for area X cell classes in subsonic and plastic song birds

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean Rate, Interbout, Hz</th>
<th>Spikewidth, ms</th>
<th>% Spikes in Burst, Singing</th>
<th>Correlation Coefficient</th>
<th>CV, ± Bout Interest</th>
<th>CV, ± Rate Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type-1</td>
<td>0.02</td>
<td>0.33</td>
<td>0.03</td>
<td>0.017</td>
<td>0.16</td>
<td>0.017</td>
</tr>
<tr>
<td>Type-2</td>
<td>0.04</td>
<td>0.13</td>
<td>0.01</td>
<td>0.021</td>
<td>0.018</td>
<td>0.021</td>
</tr>
<tr>
<td>Type-3</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Type-4</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Values are ± SD. Asterisks indicate significantly different within class comparisons of singing versus nonsinging data (P < 0.01, paired t-tests). Interclass comparisons indicate significant differences between cell classes (P < 0.01, unpaired t-tests). Peak rate was computed as 99th percentile rate during singing. For the percent of spikes within a burst, and the number of spikes per burst, bursts were defined as neural events exceeding 250 Hz (see METHODS). The correlation coefficient was measured to assess the correlation of spiking activity to syllable timing on fast timescales (see METHODS for detailed information on these parameters).

FIG. 3. Example of a type-1 neuron the activity of which was correlated with both bout onsets as well as a specific syllable. A: the raw voltage trace is plotted beneath the spectrogram (age dpf 69). Note neuronal discharge that immediately precedes the onset of the song bout. An additional spike occurs during some renditions of syllable “c.” Bottom: expanded views of the periods indicated by the blue line and red lines, respectively. B: song bout onset identified syllable raster, showing neural activity aligned to the onset of 60 consecutive bouts and the rate histogram of the raster, showing a peak in the probability of spiking that precedes bout onset. C: data are laid out as in B for neural activity aligned to 330 renditions of syllable c.

significant variability in their firing rate across repeated trials of identified syllables (Fig. 4F). Trial-to-trial cross correlations coefficients computed across smoothed spike trains over repeated syllable renditions were significantly lower than in type-1 neurons (CC = 0.017 ± 0.02, Table 1) but were still significantly different from randomly shuffled spike trains in most syllables (Fig. 4F, Table 1, n = 10/14 syllables, 6 birds, see METHODS). In the syllable-aligned rate histograms, significant peaks in firing rate were more common than rate minima, and these modulations were distributed throughout syllables (Fig. 4G).

Type-3 neurons were tonically active and did not generate high-frequency bursts

Type-3 neurons were tonically active during nonsinging periods (12.3 ± 6.9 Hz, n = 29) and exhibited increased firing rates during singing (65.1 ± 19.8 Hz, P < 0.001, paired t-test compared with nonsinging, n = 29/29 neurons; Fig. 5). This singing-related change in firing rate occurred abruptly at song onset and offset, such that the ISI distributions and the spike
train autocorrelations during interbout and during nonsinging periods differed only slightly (Fig. 5, B and C).

In contrast to all other cell types, type-3 neurons never discharged at rates >500 Hz (peak rate = 285 ± 78 Hz, Fig. 1F, Table 1). During singing, type-3 neurons had ISI distributions and autocorrelograms that resembled Poisson processes with a refractory period (CV = 1.00 ± 0.2; Fig. 5, B and C). No measure of type-3 neuronal activity was different between subsong and plastic song birds (Table 1, n = 17 and 12 neurons from subsong and plastic song birds, respectively).

In plastic song birds, type-3 neurons spiked throughout the song (sparseness index = 0.014 ± 0.010, n = 14 neurons) and showed significant rate modulations that were locked to song timing (Fig. 5F). Rate modulations in type-3 neurons were less reproducible across syllable renditions than those of type-1 neurons but more reproducible than those of type-2 neurons (Table 1, CC = 0.13 ± 0.11, pairwise cross correlations were significant in 49/55 syllables from 14 neurons in 6 birds). Significant rate increases and decreases occurred throughout syllables, and showed no clustering around syllable onsets or offsets (Fig. 5G).

**Type-4 neurons exhibit long, high-frequency bursts in excess of 1 kHz during singing**

During nonsinging periods, type-4 neurons rarely spiked, although they occasionally generated a high-frequency burst of spikes (<0.01 burst/s, n = 19/19 neurons; Fig. 6). Anecdotally, the appearance of spontaneous bursting appeared to anticipate the onset of singing within the next tens of seconds (see METHODS). Bursts in type-4 neurons were robust events, often reaching rates in excess of 1 kHz (average peak rate = 955 ± 185 Hz), and consisted of an average 6.5 ± 2.1 spike/burst. In addition, type-4 neurons rarely generated single spikes—during both singing and interbout silence, nearly all spikes were part of a burst (85 ± 9% of spikes were contained in a burst, n = 19 neurons, Table 1). Bursting was associated with a...
single peak in the ISI distributions (Fig. 6B) and a broad peak in the spike train autocorrelations at short times (Fig. 6C).

While bursts occurred more frequently during singing than during interbout silent periods (Fig. 6, D and E, singing: 6.2 ± 2.6 burst/s interbout: 2.2 ± 0.8 burst/s, P < 0.01 paired t-test compared with interbout silence), both of these states exhibited bursting far more often than in nonsinging periods, when bursts were extremely rare (<0.01 burst/s). The statistics of bursting and of singing-related changes in activity were similar in subsong and plastic song birds (Table 1, n = 10 and 9 neurons from subsong and plastic song birds, respectively).

In plastic song birds, type-4 neurons spiked throughout the song (sparseness index = 0.020 ± 0.011, n = 10 neurons) and exhibited significant variability across repeated syllable renditions (Fig. 6F; CC = 0.016 ± 0.02, 17/35 syllables exhibited significant CCs compared with shuffled spike trains, see METHODS). Despite this variability, many type-4 neurons showed significant peaks and minima in the syllable-aligned rate histograms. Peaks and minima were distributed throughout syllables and not clustered at syllable onsets or offsets (Fig. 6G).

**Salient differences in singing-related firing patterns across distinct cell types**

All four cell types had distinct singing-related firing patterns that were reflected in distinct spiking statistics. Type-1 neurons fired sparsely and occasionally generated bursts, as seen in their population-average ISI distribution and spike train autocorrelation function (Fig. 7, A and B, black trace). In addition, because the firing rate within bursts was very high relative to the mean rate, the peak of the normalized spike train autocorrelation function was nearly an order of magnitude higher than that of the other cell classes (Fig. 7B).

On superficial examination of the spike trains, neurons in the other cell classes (types 2–4) could appear to exhibit similar singing-related firing patterns: They all discharged at low rates when the bird was not singing, and all exhibited high variable firing rates during singing. However, differences among neurons types 2–4 were immediately visible in their population-average ISI distributions and spike train autocorrelations (Fig. 7, A–C). For example, type-3 neurons did not burst, exhibiting flat spike train autocorrelations (Fig. 7C) with the longest absolute refractory period (1st percentile ISI = 3.8 ± 1.2 ms, Fig. 7A). Type-4 neurons, on the other hand, discharged primarily in bursts and exhibited the shortest refractory period (1st percentile ISI = 1.1 ± 0.2 ms; Fig. 7, A–C).

To quantify the differences in firing patterns among cell types 2–4, we examined three measures of burstiness: CV of the ISI distribution, the number of spikes per burst, and the fraction of spikes within a burst. Type-3 neurons exhibited a significantly lower CV of the ISIs during singing than other cell types, and were the least bursty (Fig. 7, D and E, Table 1). In contrast, type-4 neurons exhibited the most robust bursting by all measures. In all of these measures, type-2 neurons fell intermediate to, and partially overlapping with, type-3 and -4 neurons. Note that, despite this partial overlap, type-2 neurons were clearly distinguished from the other cell types by their narrower spike width (Fig. 1).
DISCUSSION

The songbird area X is a BG nucleus specialized for singing behavior and is required for song learning (Scharff and Nottebohm 1991; Sohrabji et al. 1990). Previous studies of area X in brain slices of adult birds have identified six cell classes—two pallidal cell types and four striatal cell classes homologous to mammalian striatal neurons: MSNs, FS interneurons, cholinergic TANs, and low-threshold spiking interneurons (Farries and Perkel 2002; Farries et al. 2005). We have recorded from 185 neurons in area X in the singing bird and, based on spike waveforms and singing-related firing patterns, report six distinct cell classes. In this study, we focus on the activity of four types that exhibited low spontaneous firing rates. We hypothesize that these four cell types correspond to the four striatal cell types known to reside in area X. Because the measurement of neural activity through extracellular recording alone is insufficient to unambiguously classify neurons into distinct types, future recordings combined with juxtacellular or intracellular labeling of area X neurons will be needed to confirm these hypotheses. The two high firing classes—putative pallidal neurons—were the subject of a separate study (Goldberg et al., unpublished data).

Type-1 neurons exhibit sparse temporally precise neural activity

Several features of type-1 activity lead us to hypothesize that they are the MSNs of area X. First, type-1 neurons were distinguished by their low firing rates during singing and their sparse, temporally precise relationship to song structure (Figs. 1 and 2). The vast majority of units recorded in the striatum of primates and rodents, determined to be MSNs based on in vivo intracellular labeling studies (Wilson and Groves 1981), exhibit low firing rates and a sparse, temporally precise correlation to specific stages of a behavioral sequence (Barnes et al. 2005; Hikosaka et al. 2006; Jin et al. 2009; Kimchi et al. 2009). Second, MSNs are by far the most numerous cell type in both

**FIG. 6.** Type-4 neurons exhibit broad, high-frequency bursts in excess of 1 kHz during singing. **A:** the raw voltage trace of a representative type-4 neuron and its instantaneous firing rate (see METHODS) are plotted beneath the song spectrogram (age 64 dph). **B:** ISI distributions for the neuron shown in A during singing (blue) and interbout silence (red, see METHODS). Type-4 neurons did not spike during nonsinging periods. **C:** spike train autocorrelation for the neuron from A. Note the broad peak, resulting from the long bursts of these neurons. **D:** for each type-4 neuron, the mean firing rate during singing is plotted against the mean rate during interbout periods. **E:** the number of bursts per second during singing is plotted against burst incidence during interbout periods for each type-4 neuron. Triangles and circles, cells from subsong and plastic song birds, respectively. **F:** top: expanded view of the voltage and spectrogram from A (indicated by red bar). Inset: zoom-in on a portion of a high-frequency burst. **Bottom:** spike raster indicating spiking activity of this neuron during 80 renditions of the motif, and the rate histogram compiled from the raster plot. **G:** population histogram showing the distribution of significant rate peaks (blue) and rate minima (red) as a function of syllable timing (35 syllables from 9 type-4 neurons). Note that significant modulations occur throughout the syllables.
primates, cortical neurons that project to the striatum exhibit due to their low spontaneous firing rate (Figs. 1–3). Finally, in neurons) despite the fact that they were likely undersampled/short intervals resulting from bursts, and the peak at long intervals resulting in area X (Farries and Perkel 2002), and in the mammalian striatum (Kawaguchi et al. 1995); they were also the most frequently encountered area X neuron in our study (n = 56/115 neurons) despite the fact that they were likely undersampled due to their low spontaneous firing rate (Figs. 1–3). Finally, in primates, cortical neurons that project to the striatum exhibit firing patterns—including low firing rates with sparse activity temporally correlated with movement—that are remarkably similar to the MSNs they target (Crutcher and Alexander 1990; Turner and DeLong 2000). Similarly, projection neurons in the premotor cortical nucleus HVC (used as a common name) of adult and older juvenile birds discharge like type-1 neurons, exhibiting sparse bursts precisely locked to song timing (Kozhevnikov and Fee 2007). Thus the activity in both primate MSNs and in area X type-1 neurons is very similar to the activity in their cortical inputs.

However, a difference between the behavior of type-1 neurons in this study and MSNs recorded in mammals is that mammalian MSNs can exhibit significant variability in their mean firing rates with some neurons nearly silent (<0.1 Hz) and others significantly more active (~5 Hz) (Hikosaka et al. 2006; Wilson 1993). In contrast, type-1 neurons formed a homogenous group, silent outside of singing, and spasmodically active during singing. This may reflect the fact that area X is entirely devoted to singing behavior—and not involved in nonsinging behaviors such as grooming, eating, or flight (Feenders et al. 2008; Nottebohm et al. 1976). In contrast, in mammals, striatal circuits that serve one behavior may be spatially interwoven with those that serve another. In addition, this difference may also reflect basic interspecies differences in MSN function.

**Type-2 neurons exhibited narrow spike waveforms and brief, high-frequency bursts**

Several features of type-2 neurons lead us to hypothesize that they are the FS interneurons of area X. First, FS neurons were distinguished in brain slices from all other area X striatal cell classes by their narrow spike waveforms (Farries and Perkel 2002). This feature has been used to characterize mammalian striatal FS interneurons recorded in vivo, where juxtcortical labeling unambiguously identified striatal neurons with narrow spike waveforms as FS cells (Mallet et al. 2005). Second, during behavior, rodent FS interneurons exhibited constant discharge (~10 Hz) and were significantly less correlated with stages of a maze task than the putative MSNs (Berke 2008; Berke et al. 2004). Type-2 neurons were also constantly active (~8 Hz) and weakly correlated with song timing, exhibiting correlation coefficients significantly less than what was observed in type-1 neurons (Table 1). Finally, like type-2 neurons, striatal FS interneurons recorded in vivo exhibit brief high-frequency bursts that are associated with bimodality in their ISI distributions and peaks in their auto-correlograms (Fig. 4) (Berke 2008; Mallet et al. 2005; Schmitzer-Torbert and Redish 2004; Sharott et al. 2009).

**Type-3 neurons were tonically active and did not generate high-frequency bursts**

We hypothesize that type-3 neurons correspond to the cholinergic TANs in area X. Like TANs in the mammalian striatum, area X cholinergic neurons recorded in adult birds have large aspiny dendritic arbors, discharge spontaneously at low rates, and exhibit long spike afterhyperpolarizations (AHPs) (Farries and Perkel 2002). The long AHPs of TANs result in their having the longest absolute refractory period of all area X subtypes. Their activity in their cortical inputs.

**Fig. 7.** Distinct bursting behavior in different cell types. A: population-average ISI distributions (mean ± SE) are distinct for all neuronal types. Color code as in Fig. 1. Note the 2 peaks exhibited by type-1 neurons: the peak at short intervals resulting from bursts, and the peak at long intervals resulting from their sparse firing. B: population-average spike train autocorrelations (mean ± SE) for type-1 neurons. Right: expanded view of the peak at short times. C: population-average spike train autocorrelations (mean ± SE) for type-2-4 neurons. Note the different vertical scale. Right: expanded view of the peaks at short times. Note the relatively long refractory period of type-3 neurons. D: for each type-2, -3 and -4 neuron, the coefficient of variation of the ISI distribution (CV_{ISI}) is plotted against the fraction of all spikes that are contained within a burst. E: CV vs. the average number of spikes per burst. Left: population histogram of CV_{ISI}; bottom: population histograms for data in D and E. Color code as in Fig. 1.
spike train autocorrelograms—flat with a relatively long postspike suppression (Fig. 7D)—resemble spike train autocorrelograms of TANs recorded in primates and rodents (Raz et al. 1996; Schmitzer-Torbert and Redish 2004).

There are several features of type-3 neurons that are not perfectly consistent with TANs recorded in mammals. First, TANS recorded extracellularly in mammals exhibit wide complex spike waveforms (Apicella et al. 1991; Crutcher and DeLong 1984; Kimura et al. 1984; Raz et al. 1996), and in vivo intracellular studies demonstrate that mammalian cholinergic TANs exhibit wider action potentials than other striatal cell types (Reynolds et al. 2004; Wilson et al. 1990). In contrast, type-3 neurons did not exhibit unusually wide or complex spike waveforms (Fig. 1). Although diverging from the mammalian data, this finding is not inconsistent with intracellular recordings in area X brain slice, where many cholinergic neurons did not exhibit significantly wider spikes than other cell types (Farries and Perkel 2002).

Another difference is that the firing rates of type-3 neurons during singing (~60 Hz) were significantly higher than what is typically observed in TANS recorded in behaving mammals (~10–20 Hz) (Apicella et al. 1991; Kimura et al. 1984). While these differences could reflect interspecies variations in TAN physiology, we note that the behaviorally-related firing rates (either peak or average) of all area X neurons were higher than their mammalian counterparts by about the same factor of 3–4 (Goldberg et al. 2010).

Type-4 neurons discharged primarily in long, high-frequency bursts

The identification of type-4 neurons is made difficult by the lack of a clearly homologous firing pattern recorded in the BG of behaving mammals. To our knowledge, there are no reports of striatal neurons recorded in behaving animals that discharge almost exclusively in high-frequency bursts as we found for type-4 neurons. However, one striatal cell type—low-threshold spiking interneurons—identified in songbird area X and mammalian striatal slice recordings, exhibit pronounced calcium spikes that drive high-frequency “rebound bursts” at the offset of hyperpolarizing current injections (Farries and Perkel 2002; Kawaguchi 1993). Some area X LTS interneurons discharged spontaneously in slice with nearly all of their spikes discharging within bursts (Farries and Perkel 2002). In addition, putative LTS neurons recorded in the anesthetized rodent fire almost exclusively in bursts that reached the highest rates of all striatal cell classes (Sharott et al. 2009). Thus one possibility is that type-4 neurons correspond to the LTS interneurons in area X. An alternative possibility, however, is that type-4 neurons correspond to another cell type not yet discovered in area X slice recordings, for example low-frequency bursting units that have been reported in the primate pallidus (DeLong 1971).

Firing rate modulations of area X neurons were correlated with song temporal structure

It has been hypothesized that the BGs of both mammals and songbirds implement trial-and-error learning (Andalman and Fee 2009; Doya 2000, 1999; Kao et al. 2005; Olveczky et al. 2012).
PUTATIVE STRIATAL NEURONS IN SONGBIRD BASAL GANGLIA

In songbirds, lesions or inactivation of LMAN, the output of the AFP (Fig. 1), eliminate variability in juvenile song (Bottjer et al. 1984; Ölzeczyk et al. 2005), whereas lesions to area X result in abnormal acoustic structure and persistent variability as if exploration occurs without appropriate evaluation (Scharff and Nottebohm 1991). Thus, it has been hypothesized that area X generates “error signals” that direct song learning (Doupe et al. 2004; Doya 1999), but the nature of area X signals has remained unknown.

Here we report that, just as striatal neurons exhibit activity temporally correlated with distinct stages of a behavioral task (Aosaki et al. 1994; Barnes et al. 2005; Berke 2008), all area X cell classes exhibited temporal correlations to song on fast (~10 ms) time scales. To illustrate the population activity of area X neurons during singing, we aligned the spike trains of twenty neurons recorded in a single bird to the onset of a single syllable (Fig. 8). First, we selected nine type-1 neurons that spikedit during this syllable; other type-1 neurons (n = 6) were silent during this syllable but spiking during other syllables. Note that the type-1 neurons, putative MSNs, spiked sparsely, and preferentially at a specific time within the syllable, but different neurons recorded in the same bird spiked at different times in the song (Figs. 2 and 8). Type-2 neurons, putative FS interneurons, spiked irregularly throughout the syllable, occasionally discharging in brief, high-frequency bursts (Figs. 4 and 8). Type-3 neurons, putative TANS, were tonically active throughout the syllable, and produced regular spike trains without bursts (Figs. 5 and 8). Finally, type-4 neurons, putative LTS interneurons, discharged mostly in long, high-frequency bursts, the timing of which was weakly correlated with the song (Figs. 6 and 8). Notably, all cell types exhibited correlations to song timing that, as a population, were distributed throughout the song.

These data suggest that the population of neurons in area X can encode information continuously during singing. Such distributed coding of time during a motor sequence is useful in solving the temporal credit assignment problem and has been explicitly predicted in models where BG circuits implement reinforcement learning (Montague et al. 1996). In addition, it has recently been demonstrated that the output of the AFP generates an error signal that biases vocal output, in real time, to improve performance (Andalman and Fee 2009).

The correlations to song timing we have observed in area X play a role in generating this bias, the continuous encoding of time during singing would allow such premotor bias to be implemented at any point in the song.

Where do these correlations to song timing originate? Area X has two glutamatergic inputs. HVC neurons that project to area X exhibit brief bursts precisely time-locked to the song (Kozhevnikov and Fee 2007). In contrast, LMAN neurons spike variably with weaker correlations to song timing (Kao et al. 2005; Leonardo 2004; Ölzeczyk et al. 2005). Thus, we hypothesize that HVC plays the major role in driving timing information in area X neurons and, in particular, the type-1 neurons that strongly resemble HVC neurons in their firing patterns.

Because area X and LMAN are interconnected within a loop, it is difficult to isolate the origins of variability within the circuit. It is noteworthy that the putative MSNs (type-1) at the input layer of area X exhibit activity that is significantly more correlated to song timing than other cell types, including pallidal neurons that represent area X output (Goldberg et al. 2010). This could suggest that variability emerges within the striato-pallidal circuit. Another possibility, however, is that while type-1 neurons are driven primarily by inputs from HVC, other area X cell types that exhibit more spiking variability could be driven primarily by inputs from LMAN. Indeed, direct inputs to area X pallidal neurons from both HVC and LMAN have been demonstrated in brain slice recordings (Farries et al. 2005) and inferred in vivo (Leblois et al. 2009), and functionally resemble the “hyperdirect” pathway in mammals (Nambu et al. 2002). Thus, additional experiments such as recordings of area X neurons during LMAN inactivation are needed to resolve the issue of where variability originates in basal-ganglia thalamocortical circuits.

Summary

Our results represent the first characterization of the singing-related neural activities of multiple cell types in a BG nucleus required for song learning and provide a testable framework for the classification of striatal neurons in the singing bird. Given the highly conserved nature of vertebrate BG circuits (Reiner 2009), future studies aimed at understanding how area X cell classes interact during behavior and what role their signals may play in learning are likely to elucidate fundamental principles of BG function.

Acknowledgments

We thank Dmitry Aronov and Aaron Andalman for helpful suggestions regarding data analysis.

Grants

Funding to M. S. Fee was provided by National Institute of Deafness and Other Communication Disorder Grant R01DC-009183 and to J. H. Goldberg by the Damon Runyon Research Foundation and Charles King Trust Postdoctoral Fellowships.

References


Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw 12: 961–974, 1999.


