Origins of basal ganglia output signals in singing juvenile birds
Morgane Pidoux, Tejapratap Bollu, Tori Riccelli and Jesse H. Goldberg

You might find this additional info useful...

This article cites 62 articles, 27 of which can be accessed free at:
/content/113/3/843.full.html#ref-list-1

Updated information and services including high resolution figures, can be found at:
/content/113/3/843.full.html

Additional material and information about Journal of Neurophysiology can be found at:
http://www.the-aps.org/publications/jn

This information is current as of April 1, 2015.
Origins of basal ganglia output signals in singing juvenile birds

Morgane Pidoux, Tejapratap Bullu, Tori Riccelli, and Jesse H. Goldberg
Department of Neurobiology and Behavior, Cornell University, Ithaca, New York

Submitted 25 August 2014; accepted in final form 7 November 2014

The basal ganglia (BG) are an evolutionarily conserved set of brain nuclei critical for motor learning (Graybiel et al. 1994). BG output neurons in the internal pallidal segment (GPI) and the substantia nigra pars reticulata control behavior through their projections to the brain stem and the ventral “motor” thalamus (Albin et al. 1989; Penney and Young 1983). Across behavioral paradigms and model systems, the activities of BG output neurons are time-locked to behavior in diverse ways: neurons may exhibit rate increases or decreases in response to task cues or reward, as well as prior to, during, or following movement initiation or termination (Anderson and Horak 1985; Anderson and Turner 1991; Bryden et al. 1991; Freeze et al. 2013; Georgopoulos et al. 1983; Gulley et al. 1999; Joshua et al. 2009; Kunimatsu and Tanaka 2010; Meyer-Luehmann et al. 2002; Nambu et al. 1990; Schmidt et al. 2013; Sheth et al. 2011; Turner and Anderson 2005; 1997; Yoshida and Tanaka 2009).

The behavior-locked firing patterns of GPI neurons are controlled by multiple parallel circuits inside the BG (Fig. 1A) (Tachibana et al. 2008), including the hyperdirect pathway, as well as the direct and indirect pathways, which themselves can be modulated by striatal interneurons (Fig. 1A) (Gittis et al. 2010; Hikosaka and Wurtz 1989). Songbirds provide a unique opportunity to determine how these pathways shape behavior-locked firing of BG outputs. First, songbird area X is a striato-pallidal nucleus homologous to the mammalian BG (Fig. 1B) (Douce et al. 2005; Farries et al. 2005; Farries and Perkel 2002; Goldberg et al. 2010; Goldberg and Fee 2010). Second, during vocal babbling, BG outputs discharge with a peculiar homogeneity that, to our knowledge, has not been observed in other model systems or behaviors: GPI-like output neurons in area X exhibit short peaks in activity in the milliseconds prior to syllable onsets and brief rate decreases prior to offsets. Pre-onset decrease or pre-offset increases are never observed (Goldberg et al. 2012; Goldberg and Fee 2012). The striking homogeneity of these signals appears only during the babbling stage of vocal development. At later plastic and adult song stages, each GPI-like neuron exhibits a similar idiosyncratic timing relationship to song, increasing or decreasing its firing rate at onsets or offsets or at reproducible time points in between (Goldberg et al. 2010; Hessler and Doupe 1999; Iwasaki et al. 2013; Woolley et al. 2014).

To determine the origins of this “known” BG output signal, we recorded area X neurons during vocal babbling. We find that putative medium spiny neurons (MSNs) and fast spiking (FS) and tonically active (TAN) interneurons exhibit similar song-locked discharge as the GPI neurons: rate increases at onsets and decreases at offsets. Meanwhile, external pallidal segment (GPe)-like neurons exhibit rate decreases at onsets. These syllable locked signals were abolished by lesion of the premotor cortical nucleus HVC. Together, these findings are consistent with roles for the indirect and hyperdirect pathways in shaping BG output signals during early stages of vocal development.

MATERIALS AND METHODS

Animals. Subjects were seven control juvenile male zebra finches, 35–50 days posthatch, and four HVC-lesioned zebra finches, 40–55 days posthatch. Birds were obtained from the Cornell zebra finch breeding facility (Ithaca, NY). 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 Cornell Committee on Animal Care. All data are from juvenile birds singing undirected song.

Chronic neural recordings and histology. Recordings were carried out using a modified version of a motorized microdrive described previously (Fee and Leonardo 2001) mounted with arrays of 5–10 single electrodes (3–5 MΩ, Pd/Ir and Tungsten from microprobes.com). During microdrive implant surgery, the array was stereotactically targeted to area X based on previously established coordinates (5.8 mm anterior, 1.5 mm lateral, and 2.8 mm ventral), at a head angle 20° displaced from the flat groove on the anterior border of the bird’s skull. Electrophysiological data were band-pass filtered in custom analog circuits (0.25–15 kHz) and were acquired at 40 kHz using custom Matlab acquisition software. Units accepted for analysis had signal-to-noise ratios (average spike peak amplitude compared with SD of noise) greater than 10:1 and were sorted by threshold crossings and template matching implemented in custom Matlab code, as

Address for reprint requests and other correspondence: J. H. Goldberg, Dept. of Neurobiology and Behavior, Cornell Univ., W121 Mudd Hall, Ithaca, MA 14853 (e-mail: jesse.goldberg@cornell.edu).
remains unknown if there are different area X MSN subtypes that give rise to direct and indirect pathways via distinct outputs to GPi- and GPe-like neurons, respectively.

nidopallium (LMAN) directly to pallidal output neurons, but this pathway lacks recurrent subthalamic nucleus-GPe connections important in mammals. Second, it

activate GPi neurons through the glutamatergic subthalamic nucleus (Nambu et al. 2002). Additionally, BG output neurons may directly suppress each other’s firing through extensive collaterals (Brown et al. 2014). Finally, three classes of striatal interneuron, low-threshold spiking (LTS), fast spiking (FS) and tonically active neurons (TAN), modulate MSN firing, but their roles in shaping BG output signals remain unknown (English et al. 2012; Gittis and Kreitzer 2012; Szydlowski et al. 2013).

Fig. 1. Basal ganglia (BG) circuitry in mammals and songbirds. A: three main pathways run in parallel in mammalian BG circuits and converge on BG output neurons in the globus pallidus, internal segment (GPi). First, in the direct pathway, medium spiny neurons (MSNs) directly inhibit BG outputs via the MSN → GPi projection. Second, in the indirect pathway, the external pallidal segment (GPe), an inhibitory nucleus, is interposed between the MSNs and the GPi, such that indirect pathway MSNs may disinhibit GPi neurons through the MSN → GPe → GPi projection. Third, in the hyperdirect pathway, cortical inputs can activate GPi neurons through the glutamatergic subthalamic nucleus (Nambu et al. 2002). Additionally, BG output neurons may directly suppress each other’s firing through extensive collaterals (Brown et al. 2014). Finally, three classes of striatal interneuron, low-threshold spiking (LTS), fast spiking (FS) and tonically active neurons (TAN), modulate MSN firing, but their roles in shaping BG output signals remain unknown (English et al. 2012; Gittis and Kreitzer 2012; Szydlowski et al. 2013). B: similar circuits exist in area X nucleus of songbirds, but there are also important differences. First, no homolog of the mammalian subthalamic nucleus has yet to be reported inside area X. A hyperdirect pathway is manifest via inputs from HVC and lateral magnocellular nucleus of the anterior nidopallium (LMAN) directly to pallidal output neurons, but this pathway lacks recurrent subthalamic nucleus-GPe connections important in mammals. Second, it remains unknown if there are different area X MSN subtypes that give rise to direct and indirect pathways via distinct outputs to GPi- and GPe-like neurons, respectively.

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

\[
R(t) = \frac{1}{t_{i+1} - t_i} \text{ for } t_i < t \leq t_{i+1}
\]

where \( t_i \) is the time of the \( i \)th spike. Peak firing rates (99th percentile rate) were calculated for each neuron as the inverse of its 1st percentile interspike interval. Nonsinging firing rates were calculated from spiking activity during silent periods more than 10 s separated from singing.

Identification of cell types. BG cell types were identified on the basis of singing-related neural activity [for putative GPe, GPi, MSN, TAN, and low-threshold spike (LTS)] and spike waveform (for putative FS cells) as described previously (Goldberg et al. 2010; Goldberg and Fee 2010). For each neuron, the spike width was calculated as the half-width of the average of 50 spike waveform examples. Units with spike widths less than 0.06 ms were part of a distinct cluster and were identified as putative FS neurons (Goldberg and Fee 2010). Neurons with mean firing rate during singing less than 5 Hz and spike width greater than 0.06 ms were classified as putative MSNs; neurons with mean firing rate during singing greater than 10 Hz and peak firing rate (inverse of the 1st percentile interspike interval) less than 600 Hz were identified as putative TANs; those with peak firing rate greater than 600 Hz were identified as LTS neurons. Pallidal neurons were distinguished from the above striatal cell types by mean firing rate during singing greater than 50 Hz. Furthermore, to distinguish GPe neurons from GPi pallidal neurons, we used previously established criteria that leverage the distinct spectral properties of GPe and GPi firing: GPe-like neurons exhibit slow rate modulations with long burst and pauses, whereas GPi-like neurons exhibit high-frequency modulations without long bursts or pauses. To quantitatively distinguish these neurons, we smoothed the IFR with an finite impulse response equiripple filter with a pass band below 25 Hz and stop band greater than 75 Hz, with 80-dB attenuation in the stop band. To remove the DC offset (high average firing rate), IFRs were mean-subtracted using an infinite impulse response 1-Hz high-pass filter generated with coefficients \( b = [1 \ -1] \) and \( a = [1 \ -0.9988] \). The peak of this smoothed IFR produces two distinct clusters that separate GPe-like neurons, which exhibit peaks in smoothed IFR greater than 250 Hz, from GPi-like neurons, which exhibit peaks less than 250 Hz (Goldberg et al. 2010).

Analysis of correlations of neural activity to syllable onsets and offsets. Statistical significance of rate increases and decreases in syllable onset (or offset) aligned histograms was calculated as described previously (Goldberg and Fee 2012). Neural activity was aligned to the 200 ms preceding and following all syllable onsets and offsets. Only cells recorded for greater than 50 syllables were included in the study. To determine the significance of firing-rate peaks and troughs, a rate histogram (bin size, 10 ms) was generated of the real data. Then a surrogate histogram was generated in which each trial of syllable-aligned neural activity was time-shifted by a uniformly distributed random amount over a range equal to the duration of the histogram (400 ms). The shift was circular, such that spikes wrapped around to the beginning of the histogram, preserving the overall spike
statistics of the data. The minimum and maximum of the surrogate rate histogram was then obtained with 1,000 repetitions of randomly shifted data. P values for the rate minimum and maximum of the real data set were calculated by analyzing the frequency with which shifted data sets generated larger maxima or smaller minima in firing rates. Maximal rate increases and decreases with P values < 0.05 were considered significant. The amplitude of the rate modulation was computed as the maximal deviation from the average rate during the ±200 ms around syllable onset (or offset). The timing of the rate change relative to the syllable onset (or offset) was computed as the time of maximal rate change above (for rate increases) or below (for rate decreases) the baseline rate. To construct population rate histograms (as in Fig. 2D), histograms from individual cells were mean-subtracted and then averaged across the group.

Most syllables during vocal babbling are less than 200 ms in duration, and, because syllable onsets are preceded by syllable offsets and syllable offsets are followed by syllable onsets, average rate histograms of activity ±200 ms around a syllable boundary include firing rate changes, not only at the zero time point of the boundary of interest, but also over the preceding and following onsets and offsets (Aronov et al. 2008). For this reason, rate histograms of neurons with onset- and offset-locked firing appear periodic and exhibit side peaks and troughs. Thus, for each histogram with a significant rate change, we defined the “dominant” rate change (either an increase or a decrease) as the one that exhibited the largest deviation from the neuron’s average firing rate (Table 1). For each dominant rate change, we computed five values: 1) its magnitude (in spike rate, Hz); 2–4) the timing of its peak (2), onset (3), and offset (4) (in milliseconds relative to syllable boundary, as in Fig. 2, E and J, and reported in Table 1); and 5) its modulation width (defined as the duration between its offset and onset). The onset was defined as the bin that exceeded 2 SDs from the mean in the baseline rate. The offset was the bin that returned to below 2 SDs from mean. The modulation width (ms) was the time separating the onset and offset of the rate change (Table 1).

For MSNs and putative FS cells with very low baseline firing rates and strong peaks at syllable onsets, significant peaks were also observed in offset-aligned histograms; closer inspection of offset-aligned rasters showed that these offset-peaks were due to high probability of firing in relation to the following (e.g., see Fig. 5D) or preceding (e.g., see Fig. 5F) syllable onset. Thus, for MSNs and FS interneurons, we also defined a dominant activation as occurring at either onset or offset, the dominant one being the one with the larger peak. In all cases, this dominant rate change occurred in the syllable onset aligned histograms. Note that, because these cell classes exhibit such low baseline firing rates, we did not test for significance of rate decreases in syllable-aligned histograms of MSN or FS activity.

To quantitatively test if average firing rate peaks (and dips) were clustered at syllable boundaries beyond predicted by chance, we determined, for each neuron, the timing of the dominant rate change at onset and offset, as described above. Next, for each cell class, we constructed a “real data” histogram that plotted the number of cells exhibiting a dominant rate change as a function of time relative to the syllable boundary (−0.3 to +0.3 s relative to syllable boundary; 0.2-s bins). We also computed 1,000 surrogate histograms with the same data, but with each cell’s timing randomly shuffled. We computed the P value of the peak observed in the real data by analyzing the frequency with which shuffled data sets generated a larger peak than what was observed. P values < 0.01 were considered significant.

To determine the reliability of neuronal firing rate changes across syllables, we computed the percentage of syllables in which a neuron exhibited an increased rate at the time of its dominant rate change (quantified in the average syllable onset aligned rate histogram). Specifically, for each neuron with a dominant rate change, we examined the mean IFR that occurred in a 60-ms time window centered at the time of its dominant rate change and then determined the frequency with which this mean was greater than that observed in the
HVC. These neurons were confirmed to be absent following HVC birds, tracer injection into area X results in pronounced labeling of lesions, as previously described in detail (Aronov et al. 2008; Goldberg and Fee 2010; Woolley et al. 2014). MSNs fired similarly sparsely in babbling birds, exhibiting a small but significant increase in firing rate during singing (mean rate during silent periods: 0.22 ± 0.8 Hz vs. mean rate during singing 0.80 ± 0.5 Hz, P < 0.05, paired t-test, n = 20 neurons) and discharging during only 10.6 ± 6.0% of syllables (mean ± SD). Inspection of syllable-locked firing revealed that most exhibited a significantly higher probability of spiking at syllable onsets (rate increase of 4.3 ± 1.7 Hz occurring at −9.0 ± 32 ms relative to syllable onsets, statistically significant rate increases in n = 15/20 neurons, Table 1, Fig. 3F, see MATERIALS AND METHODS). MSNs with onset peaks also exhibited peaks in offset-aligned histograms. However, examination of syllable offset-aligned spike rasters that were sorted by the arrival time of the preceding or following syllable onset clearly showed that these peaks were attributable to the discharge related to the neighboring onsets (e.g., see Fig. 3, H and I). Thus for each MSN we defined the dominant peak as occurring in either onset or offset-aligned histograms as the one with the larger rate change from baseline (see MATERIALS AND METHODS). Of the 15/20 neurons that exhibited statistically significant syllable-locked activations, all exhibited dominant peaks at syllable onsets, and none of them at offsets (Fig. 3, Table 1, peaks were clustered at offsets, P < 0.001, see MATERIALS AND METHODS).

The timing of MSN firing does not support a direct-pathway origin of GPe onset peaks. If anything, MSN-GPi inhibition would produce pauses prior to syllable onsets and peaks prior to syllable offset, the opposite of what was observed in the GPe-like neurons. Yet MSNs also target GPe-like neurons, which can suppress firing of GPe neurons. If this pathway contributes to syllable timing signals in GPe neurons, one would expect GPe neurons to exhibit rate decreases prior to onsets, and increases prior to offsets. To test this hypothesis, we recorded GPe-like neurons in area X. As reported previously, these neurons discharged like many primates GPe neu-

### Table 1. The timing and magnitude of syllable-locked firing patterns across putatively identified cell classes during vocal babbling

<table>
<thead>
<tr>
<th></th>
<th>GPI</th>
<th>MSN</th>
<th>GPe</th>
<th>FS</th>
<th>TAN</th>
<th>LTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant rate change, syllable onset, Hz</td>
<td>18.9 ± 6.1</td>
<td>4.3 ± 1.7</td>
<td>−19.8 ± 7.4</td>
<td>15.9 ± 14.0</td>
<td>20.0 ± 8.2</td>
<td>N/S</td>
</tr>
<tr>
<td>Peak of dominant rate change, syllable onset, ms</td>
<td>−16 ± 32</td>
<td>−9 ± 32</td>
<td>−18 ± 30</td>
<td>−8 ± 10</td>
<td>−15 ± 13</td>
<td>N/S</td>
</tr>
<tr>
<td>Onset of dominant rate change, syllable onset, ms</td>
<td>−45 ± 20</td>
<td>−3 ± 60</td>
<td>−37 ± 30</td>
<td>−27 ± 17</td>
<td>−43 ± 13</td>
<td>N/S</td>
</tr>
<tr>
<td>Width of dominant rate change, onset, ms</td>
<td>41 ± 20</td>
<td>30 ± 13*†</td>
<td>47 ± 11</td>
<td>38 ± 13</td>
<td>53 ± 14</td>
<td>N/S</td>
</tr>
<tr>
<td>No. cells with dominant rate increase at onset</td>
<td>22/26</td>
<td>15/20</td>
<td>0/15</td>
<td>4/5</td>
<td>11/13</td>
<td>0/10</td>
</tr>
<tr>
<td>No. cells with dominant rate decrease at onset</td>
<td>0/26</td>
<td>N/A</td>
<td>12/15</td>
<td>N/A</td>
<td>0/13</td>
<td>0/10</td>
</tr>
<tr>
<td>Onset reliability, %syllable onsets exhibiting dominant rate change</td>
<td>60 ± 6</td>
<td>N/A</td>
<td>59 ± 5</td>
<td>N/A</td>
<td>67 ± 7*‡</td>
<td>N/S</td>
</tr>
<tr>
<td>Dominant rate change, syllable offset, Hz</td>
<td>−13.7 ± 6.6</td>
<td>N/A</td>
<td>15.0 ± 5.0</td>
<td>N/A</td>
<td>−17.2 ± 12.7</td>
<td>N/S</td>
</tr>
<tr>
<td>Peak of dominant rate change, syllable offset, ms</td>
<td>−29 ± 42</td>
<td>N/A</td>
<td>−1 ± 60</td>
<td>N/A</td>
<td>−17 ± 16</td>
<td>N/S</td>
</tr>
<tr>
<td>Onset of dominant rate change, syllable offset, ms</td>
<td>−43 ± 46</td>
<td>N/A</td>
<td>−27 ± 67</td>
<td>N/A</td>
<td>−43 ± 13</td>
<td>N/S</td>
</tr>
<tr>
<td>Width of dominant rate change, offset, ms</td>
<td>37 ± 14</td>
<td>N/A</td>
<td>38 ± 16</td>
<td>N/A</td>
<td>−53 ± 13</td>
<td>N/S</td>
</tr>
<tr>
<td>No. cells with dominant rate increase at offset</td>
<td>40/26</td>
<td>0/20</td>
<td>6/15</td>
<td>0/5</td>
<td>0/13</td>
<td>0/10</td>
</tr>
<tr>
<td>No. cells with dominant rate decrease at offset</td>
<td>14/26</td>
<td>N/A</td>
<td>0/15</td>
<td>N/A</td>
<td>7/13</td>
<td>0/10</td>
</tr>
<tr>
<td>Offset reliability, %syllable offsets exhibiting dominant rate change</td>
<td>55 ± 15</td>
<td>N/A</td>
<td>55 ± 17</td>
<td>N/A</td>
<td>56 ± 3</td>
<td>N/S</td>
</tr>
</tbody>
</table>

Values are means ± SD across cells within a cell class. GPI, globus pallidus internal segment; MSN, medium spiny neurons; GPe, globus pallidus external segment; FS, fast spiking; TAN, tonically active; LTS, low-threshold spiking; N/A, not applicable. For each cell class, the magnitude of the dominant maximal rate change (Hz), its timing relative to syllable onset and offset, its reliability across syllables, and the number of cells with statistically significant dominant rate increases and decreases in syllable onset- and offset-aligned histograms are shown (see MATERIALS AND METHODS). Note that these data were acquired from syllable-locked rate histograms that exhibited statistically significant rate changes. N/S indicates that there were no such histograms in that cell class dataset. Note also that rate decreases in rate histograms on onset/offset reliability were not computed in MSN and FS cell classes due to their sparse firing. A different metric, the percentage of syllables with a spike, is shown in Table 3. See MATERIALS AND METHODS for analysis details. *P < 0.05 in unpaired t-test compared with the cell classes indicated. †<GPe and TAN. ‡>GPe and GPe.
Three types of striatal interneuron are conserved between songbirds and mammals, FS, LTS and TAN, yet it remains unclear how their activities shape MSN and GP firing. We recorded units putatively identified as these subclasses and found that both FS and TAN cell types also exhibited rate peaks at syllable onsets and decreases at syllable offsets (Figs. 5 and 6, Table 1). Meanwhile, LTS neurons exhibited highly randomly timed high-frequency bursts, with weak, if any, relationship to syllable timing (Fig. 7, Table 1).

Note that all of the syllable-locked signals reported thus far are computed from syllable-aligned rate histograms, which reflect the average firing of a neuron across syllables. Yet a neuron need not increase its discharge at every syllable to yield a strong peak in the histogram. For each neuron in the GPi, GPe and TAN cell classes, we defined its “reliability” as the percentage of syllables during which an observed rate change occurred in the direction of the dominant rate change (see MATERIALS AND METHODS). At syllable onsets, TAN neurons exhibited increased firing rates prior to 67 ± 7% of syllables (P < 0.01 compared with GPi: 60 ± 6%; and to GPe: 59 ± 5%, Table 1). There were no significant intergroup differences in reliability of syllable-offset locked rate changes (Table 1).

We next wondered what input to area X might cause these remarkably homogenous firing patterns. The two main cortical inputs to area X, HVC and lateral magnocellular nucleus of the anterior nidopallium (LMAN), play different roles in vocal babbling. LMAN lesions abolish singing altogether in juvenile babbling birds, and abolish vocal variability in older plastic song birds, causing premature song crystallization (Aronov et al. 2008). Thus HVC lesions during vocal babbling (Aronov et al. 2008). Thus HVC lesions provide an opportunity to record area X neurons during a singing behavior that is highly similar to vocal babbling.

To test how syllable onset and offset aligned firing patterns depend on inputs from HVC, we recorded GPi, GPe and MSN cell classes after complete bilateral HVC lesion in juvenile birds (see MATERIALS AND METHODS). As reported previously, song after HVC lesion resembled vocal babbling (Figs. 8 –10) (Aronov et al. 2008; Chen et al. 2014b; Goldberg and Fee 2011). We also found that the basic spiking statistics of pallidal neurons were not significantly affected by HVC lesion, and the same analyses that distinguished these cell types in intact birds distinguished them in HVC-lesioned birds (Table 2). In MSNs, average firing rate during singing, but not during silent periods, was slightly but significantly elevated in HVC-lesioned birds (average rate during singing, control: 0.88 ± 0.50 Hz vs.

Fig. 3. Putative MSNs in area X exhibit rate peaks at syllable onsets during vocal babbling. A: spectrogram showing vocal babbling of a juvenile bird is shown above the spiking activity of a single MSN in area X. B: expanded view of song (top) and spiking activity (bottom) aligned to the onset of a single syllable (marked with blue asterisks). C: syllable onset aligned raster plot showing the spiking activity of this neuron during 1,800 syllable renditions. The raster rows are sorted by the duration to the syllable offset (curved blue line at right of raster). To visualize sparse spikes as raster tics, tic height was 5 milliseconds. D and E: syllable onset-aligned rate histograms for the neuron in shown in A–C (D) and for all MSNs recorded in area X (E) (shading indicates ± SE) (n = 20). F: scatter plot showing the magnitude of significant dominant rate increases (black triangles) plotted against the time, relative to syllable onset, at which they occurred. Note that the x-axis is the same for B–F. G–K: data are plotted exactly as in B–E for syllable offsets. In the raster in H, the data are sorted by the arrival time of the following syllable onset. Note that the post-offset spiking is aligned to the following syllable onset. K: 0/20 MSNs exhibited their dominant rate peak at syllable offsets.
Seven percent of MSNs in area X exhibited rate peaks in syllable onset histograms (Fig. 9). Only 2/19 GPe-like neurons had rate increases at syllable onsets, and none had significant rate decreases at offsets. Together, these data suggest that HVC is a principal driver of syllable timing-related firing in area X of babbling birds.

The importance of HVC in driving syllable-locked firing raises the specific prediction that timing signals in area X should be syllable specific. Specifically, in subsong and early plastic songbirds, short syllables less than 50 ms are abolished by LMAN lesion, elongated by mild LMAN cooling, but not strongly affected by HVC lesion or cooling (Aronov et al. 2011; Veit et al. 2011). To test if area X neurons exhibit diminished timing-related signals at short syllables, we examined onset- and offset-locked firing at short (<50 ms) syllables. The magnitude of dominant peaks and dips at short-syllable onsets was reduced in all cell types (Table 4), consistent with a diminished role of HVC in contributing to the initiation of brief syllables.

**DISCUSSION**

We studied the origins of BG output signals during vocal babbling in juvenile zebra finches. Previous studies showed that, at this early stage of vocal development, the GPe-like BG output neurons in the songbird area X exhibit rate increases prior to syllable onsets and decreases prior to syllable offsets. Here we report that syllable-locked firing is similarly homogenous for upstream BG cell classes. Putative MSN, TAN, and FS cell classes exhibited transient rate increases at syllable onsets, and rate decreases at offsets, while GPe-like neurons exhibited transient decreases at onsets. Meanwhile, putative LTS neurons were unmodulated by syllable timing. HVC lesions did not significantly affect babbling behavior but largely abolished these song timing signals inside area X.

Together, these findings are consistent with a role for the MSN-GPe-GPi indirect pathway inside area X in transforming cortically derived timing signals into a BG output signal. Yet several caveats are relevant to this interpretation, and other scenarios cannot be ruled out. First, HVC projects directly onto pallial neurons in area X (Faries et al. 2005). Thus the peak in GPe firing at syllable onsets might also arise from hyperdirect inputs form HVC (Fig. 1B). Support for a role of the hyperdirect pathway is that many MSNs exhibited onset peaks (Fig. 3F) after the occurrence of GPe onset troughs (Fig. 4E) and GPe onset peaks (Fig. 2E; Table 1). These “late-peak” MSN neurons are therefore unlikely to play a role in sculpting the average BG output signal and may reflect more complex processing inside the BG, perhaps related to syllable phonology or elongation. Note that, for this hyperdirect pathway alone to explain the syllable locked firing of the other neurons in area X, GPe-GPe-MSN feedback would be required, which, although not impossible, is opposite the direction of signal propagation described in this circuit in mammals (Hikosaka 2007; Jaeger and Kita 2011; Nambu et al. 2002).

A second caveat is that we were unable to determine in our recordings which “type” of MSN we recorded and thus have no direct evidence for their participation in direct or indirect pathways or both. Specifically, area X MSNs may express exclusively D1 receptors or D2 receptors, but, unlike mammalian MSNs, most songbird ones express both (Kubikova et al. 2010). Based on evolutionarily conserved patterns (Stephenson-Jones et al. 2012), it is possible that these distinct subtypes project differentially to GPe- and GPe-like neurons in area X, but this remains unknown. Assuming that we randomly sampled MSN subtypes in area X, it is likely that the 20 we recorded represented a mix of MSN cell types. If this is the case, then this would suggest that both direct and indirect pathway MSNs synchronously exhibit rate increases at syllable initiation, similar to what was recently observed in mice in the context of locomotor initiation (Cui et al. 2013). However, we cannot rule out the possibility that the 5/20 MSNs that did not exhibit strong syllable-locked firing were a specific MSN subtype.

A third caveat is that, while GPe firing increases, on average, prior to syllable onsets, rate increases were not observed prior to each individual syllable (see reliability in Table 1) (Goldberg and Fee 2012). By focusing on average syllable locked rate histograms, we did not emphasize neural firing that varies on a syllable-by-syllable basis; instead, we focused on possible sources of the robust GPe onset signal only observable by averaging across syllables. Thus, while our data are consistent with a role for the indirect and hyperdirect pathways in shaping

---

**Table 2. Cell-class specific spiking characteristics of pallidal neurons persist following HVC lesion**

<table>
<thead>
<tr>
<th></th>
<th>GPI</th>
<th>GPI (No HVC)</th>
<th>GPe</th>
<th>GPe (No HVC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate, nonsinging, Hz</td>
<td>119.5 ± 48.0</td>
<td>122.1 ± 27.8</td>
<td>113.5 ± 28.8</td>
<td>108.2 ± 37.6</td>
</tr>
<tr>
<td>Rate, singing, Hz</td>
<td>263.1 ± 67.5</td>
<td>249.2 ± 58.3</td>
<td>125.5 ± 33.3</td>
<td>152.3 ± 49.3</td>
</tr>
<tr>
<td>CV, nonsinging</td>
<td>0.33 ± 0.13</td>
<td>0.30 ± 0.09</td>
<td>0.36 ± 0.08</td>
<td>0.36 ± 0.10</td>
</tr>
<tr>
<td>CV, singing</td>
<td>0.50 ± 0.10</td>
<td>0.43 ± 0.09</td>
<td>0.87 ± 0.27</td>
<td>0.85 ± 0.35</td>
</tr>
<tr>
<td>99th percentile ISI, nonsinging, ms</td>
<td>16.4 ± 6.8</td>
<td>15.5 ± 6.7</td>
<td>20.2 ± 7.8</td>
<td>22.9 ± 11.7</td>
</tr>
<tr>
<td>99th percentile ISI, singing, ms</td>
<td>11.2 ± 4.4</td>
<td>12.1 ± 5.6</td>
<td>42.5 ± 30.0</td>
<td>34.9 ± 22.5</td>
</tr>
<tr>
<td>Peak in smoothed firing rate (range: min–max)</td>
<td>195 ± 21 (153–228)</td>
<td>189 ± 28 (151–249)</td>
<td>404 ± 88 (266–583)</td>
<td>404 ± 112 (253–587)</td>
</tr>
</tbody>
</table>

Values are means ± SD across cells within a cell class. Singing-related changes in firing rates and interspike interval (ISI) statistics were not affected by HVC lesion. The peak in the smoothed instantaneous firing rate, the metric used to distinguish GPe- from GPe-like neurons, was also not affected (see MATERIALS AND METHODS). CV, coefficient of variation.
This average BG output signal, on a trial-by-trial basis, signal propagation is likely to be more complex.

A final caveat relates to our assignment of cell class type in our area X recordings. In past work, we discovered six distinct singing-related firing patterns in area X (Goldberg et al. 2010; Goldberg and Fee 2010), and we attempted to map their identities to the six cell classes that were characterized in area X in vitro (Carrillo and Doupe 2004; Farries et al. 2005; Perkel et al. 2002). This mapping was based on the similarities to well-characterized firing patterns of potentially homologous mammalian BG cell types. Yet, because extracellular recording alone is not sufficient to unambiguously identify a neuron type, our cell type assignments in this paper represent our current best hypotheses for cell identity. This caveat is particularly important for the striatal interneurons whose firing patterns during behavior remain poorly understood.

An outstanding question is the functional significance of the syllable timing signals inside area X during babbling. The production of vocal babbling is not strongly affected by lesions of HVC or of area X (Aronov et al. 2008, 2011; Chen et al. 2014a; Goldberg and Fee 2011); therefore, these signals are not necessary for vocal motor production and are unlikely to be actually driving syllable onsets and offsets. Instead, because area X lesions block learning (Goldberg and Fee 2011; Scharff and Nottebohm 1991), syllable timing signals in area X might reflect a learning signal important for advancing the subsong-plastic song transition. Specifically, this transition is the earliest manifestation of vocal motor learning in birdsong and is characterized by the emergence of a “protosyllable” of discernable phonological and temporal structure that is repeated in a rhythmic fashion (Aronov et al. 2011; Tchernichovski et al. 2001). The critical neural event underlying protosyllable emergence is the linkage of a syllable offset to its own onset. This linkage enables a syllable to become stereotyped in duration and to be repetitively produced: two events that underlie the first emergence of song rhythm (Aronov et al. 2011; Lipkind et al. 2013; Saar and Mitra 2008; Veit et al. 2011). Given the importance of syllable onset and offset times in this process, an intriguing hypothesis is that syllable boundary signals observed inside area X in this study are important for the early emergence of temporal structure in the song itself. To test this and other hypotheses, it will be necessary to manipulate activity in area X during babbling, for example, with opto-chemogenetics. Manipulations that influence song in “real time” would support a model where area X is an “actor” that biases vocalizations. Manipulations that only affect protosyllable learning would support a model where area X may function more like a “critic” (Doya and Sejnowski 1995; Fee and Goldberg 2011; Sutton and Barto 1998). Notably, protosyllables resemble “reduplication” events that occur early in human babbling (Doupe and Kuhl 1999; Oller 1980). Because BG disorders impair vocal learning in humans (Enard 2011), determining the functional significance of BG timing signals in babbling birds may illuminate neural mechanisms by which the earliest exploratory utterances acquire nascent temporal patterns during learning.

GRANTS

Funding to J. H. Goldberg was provided by the National Institute of Neurological Disorders and Stroke (Grant R00-NS-067062), the Klingenstein Neuroscience Foundation, and the Pew Biomedical Scholars Award.
No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES


Brown J, Pan WX, Dudman JT. The inhibitory microcircuit of the substantia nigra provides feedback gain control of the basal ganglia output. Elife 3: e02397, 2014.


Farriss MA, Ding L, Perkel DJ. Evidence for “direct” and “indirect” pathways through the song system basal ganglia. J Comp Neurol 484: 93–104, 2005.


Fig. 6. Putative TAN interneurons in area X exhibit rate increases at syllable onsets during vocal babbling. A: spectrogram showing vocal babbling of a juvenile bird is shown above the spiking activity (middle) and IFR (bottom) of a putative TAN interneuron recorded in area X. B: expanded view of song (top) and spiking activity (bottom) aligned to the onset of a single syllable (marked with blue asterisks). C: syllable onset aligned raster plot showing the spiking activity of this neuron during 1,800 syllable renditions. The data are sorted by the duration to the syllable offset (curved blue line at right of raster). Tic height in the raster is 1 syllable. Syllable-onset-aligned rate histograms for the neuron shown in A–C (D) and for all TAN interneurons recorded in area X (shading indicates ± SE) (n = 13) (E). F: scatter plot showing the magnitude of significant rate increases (black triangles) and rate decreases (red inverted triangles) plotted against the time, relative to syllable onset, at which they occurred. Note that the x-axis is the same for B–F: G–K: data are plotted exactly as in B–E for syllable offsets. H: the syllable-offset aligned raster is sorted by the arrival time of the following syllable onset (curved blue line at right of raster).
Fig. 7. Putative LTS interneurons in area X do not exhibit strong syllable-locked firing during vocal babbling. A: spectrogram showing vocal babbling of a juvenile bird is shown above the spiking activity (middle) and IFR (bottom) of a putative LTS interneuron recorded in area X. B: expanded view of song (top) and spiking activity (bottom) aligned to the onset of a single syllable (marked with blue asterisks). C: syllable onset aligned raster plot showing the spiking activity of this neuron during 1,100 syllable renditions. The data are sorted by the duration to the syllable offset (curved blue line at right of raster). Tic height in the raster is 4 syllables. Syllable onset-aligned rate histograms for the neuron shown in A–C (D) and for all LTS interneurons recorded in area X (shading indicates ± SE) (n = 10) (E). Note that the x-axis is the same for B–E. F–I: data are plotted exactly as in B–E for syllable offsets. H: the syllable-offset aligned raster is sorted by the arrival time of the following syllable onset (curved blue line at right of raster).
Fig. 8. HVC lesions abolish syllable-locked firing in area X GPi neurons. A–D: histological verification of complete HVC lesion. A and B: in unlesioned birds, HVC is clearly visible in dark-field (A) and via fluorescent cell bodies retrogradely labeled from injection of Alexa dextran tracer into area X (white dotted line, see MATERIALS AND METHODS) (B). C and D: electrolytic lesion of HVC was confirmed by visible destruction of the structure visible in dark-field (C) and in the failure to observe retrogradely labeled neurons following injection of tracer into area X (D). Scale bar in A applies to A–D. E: spectrogram showing post-HVC lesion song of the juvenile bird whose brain is shown in C and D. This song is shown above the IFR of a GPi neuron in area X. F: expanded view of song (top), spiking activity (middle) and IFR (bottom) aligned to the onset of a single syllable (marked with blue asterisks). Syllable onset-aligned rate histograms for the neuron are shown in E and F (G) and for all GPi-like neurons recorded in area X after HVC lesion (shading indicates ± SE) (n = 17) (H). Note that the x-axis is the same for F–H, I–K: data are plotted exactly as in F–H for syllable offsets. In G and H and J and K, the ordinate scale was matched to the scale used in homologous plots for control GPi neurons in Fig. 2.
Fig. 9. HVC lesions abolish syllable-locked firing in area X GPe neurons. A: spectrogram showing post-HVC lesion song of a juvenile bird is plotted above the spiking activity (middle) and IFR (bottom) of a single GPe-like neuron recorded in area X of a bird with complete bilateral HVC lesion. B: expanded view of song (top) and spiking activity (bottom) aligned to the onset of a single syllable (marked with blue asterisks). C: syllable onset aligned raster plot showing the spiking activity of this neuron during 400 syllable renditions. The data are sorted by the duration to the syllable offset (curved blue line at right of raster). Tic height in the raster is 1 syllable. Syllable onset-aligned rate histograms for the neuron in shown in A–C (D) and for all GPe neurons recorded in area X of HVC-lesioned birds (shading indicates /H11006 SE) (E). F–I: data are plotted exactly as in B–E for syllable offsets.

Fig. 10. HVC lesions abolish syllable-locked firing in area X MSNs. A: spectrogram showing post-HVC lesion song of a juvenile bird is shown above the spiking activity of a single MSN in area X. B: expanded view of song (top) and spiking activity (bottom) aligned to the onset of a single syllable (marked with blue asterisks). C: syllable onset aligned raster plot showing the spiking activity of this neuron during 500 syllable renditions. The data are sorted by the duration to the syllable offset (curved blue line at right of raster). Tic height in the raster was 5 syllables. D and E: syllable onset-aligned rate histograms for the neuron in shown in A–C (D) and for all MSNs recorded in area X (E) (shading indicates ± SE) (n = 20). F–H: data are plotted exactly as in B–E for syllable offsets. H: in the raster, the data are sorted by the arrival time of the following syllable onset.

Table 3. Spiking characteristics of putative MSNs following HVC lesion

<table>
<thead>
<tr>
<th></th>
<th>MSN</th>
<th>MSN (No HVC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Rate, nonsinging, Hz</td>
<td>0.22 ± 0.79</td>
<td>0.64 ± 0.72</td>
</tr>
<tr>
<td>Rate, singing, Hz</td>
<td>0.88 ± 0.50</td>
<td>1.81 ± 1.18*</td>
</tr>
<tr>
<td>Syllables with spike, %</td>
<td>10.6 ± 6.0</td>
<td>18.1 ± 7.7*</td>
</tr>
</tbody>
</table>

Values are means ± SD across cells within a cell class; n, no. of medium spiny neurons (MSNs). *P < 0.05 in Wilcoxon rank sum tests.
Table 4.  *Syllable-locked rate modulations during syllables <50 ms in duration*

<table>
<thead>
<tr>
<th></th>
<th>GPI</th>
<th>MSN</th>
<th>GPe</th>
<th>FS</th>
<th>TAN</th>
<th>LTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant rate change, short syllable onsets, Hz</td>
<td>11.3 ± 15.0*</td>
<td>0.9 ± 2.4*</td>
<td>−10.8 ± 10.0*</td>
<td>4.8 ± 6.9</td>
<td>10.8 ± 14.0*</td>
<td>N/S</td>
</tr>
<tr>
<td>Dominant rate change, short syllable offsets, Hz</td>
<td>−14.2 ± 19.0</td>
<td>N/A</td>
<td>14.4 ± 22.0</td>
<td>N/A</td>
<td>−8.6 ± 6.8</td>
<td>N/S</td>
</tr>
</tbody>
</table>

Values are means ± SD across cells within a cell class. Data are plotted as in Table 1 showing magnitude of the dominant rate change for short syllables. *P < 0.05, indicates that the rate change was significantly less than that observed for all-duration syllables, shown in Table 1.