Role of the Zebra Finch Auditory Thalamus in Generating Complex Representations for Natural Sounds

Noopur Amin, Patrick Gill and Frédéric E. Theunissen

J Neurophysiol 104:784-798, 2010. First published 16 June 2010; doi:10.1152/jn.00128.2010

J Neurophysiol 104: 784–798, 2010.
First published June 16, 2010; doi:10.1152/jn.00128.2010.

Role of the Zebra Finch Auditory Thalamus in Generating Complex Representations for Natural Sounds

Noopur Amin,[1] Patrick Gill,[3] and Frédéric E. Theunissen[1,2,3]

[1]*Helen Wills Neuroscience Institute and* [2]*Psychology Department, University of California, Berkeley, California; and* [3]*School of Electrical and Computer Engineering, Cornell University, Ithaca, New York*

Submitted 29 January 2010; accepted in final form 10 June 2010

Amin N, Gill P, Theunissen FE. Role of the zebra finch auditory thalamus in generating complex representations for natural sounds. *J Neurophysiol* 104: 784–798, 2010. First published June 16, 2010; doi:10.1152/jn.00128.2010. We estimated the spectrotemporal receptive fields of neurons in the songbird auditory thalamus, nucleus ovoidalis, and compared the neural representation of complex sounds in the auditory thalamus to those found in the upstream auditory midbrain nucleus, mesencephalicus lateralis dorsalis (MLd), and the downstream auditory pallial region, field L. Our data refute the idea that the primary sensory thalamus acts as a simple, relay nucleus: we find that the auditory thalamic receptive fields obtained in response to song are more complex than the ones found in the midbrain. Moreover, we find that linear tuning diversity and complexity in ovoidalis (Ov) are closer to those found in field L than in MLd. We also find prevalent tuning to intermediate spectral and temporal modulations, a feature that is unique to Ov. Thus even a feed-forward model of the sensory processing chain, where neural responses in the sensory thalamus reveals intermediate response properties between those in the sensory periphery and those in the primary sensory cortex, is inadequate in describing the tuning found in Ov. Based on these results, we believe that the auditory thalamic circuitry plays an important role in generating novel complex representations for specific features found in natural sounds.

INTRODUCTION

One of the key functions of the auditory system is to recognize and categorize vocalizations in terms of their behavioral relevance. Much of our understanding on how this function is achieved relies on a feed forward hierarchical model in which auditory neurons are tuned to more complex features as the auditory signal traverses up the processing stream. This model is well supported by experimental evidence. At the auditory periphery, single neurons are tuned to simple spectral and temporal features (e.g., Feng et al. 1975; Kemp 1979), whereas, at the other extreme, network dynamics in auditory cortical regions reflect more complex cognitive processing such as categorical perception of sound classes (Ohl et al. 2001). At intermediate levels, selectivity for sound features found in vocalizations has been reported in neurons in the auditory midbrain and forebrain (or pallium) of both mammals (Liu and Schreiner 2007; Newman and Wollberg 1978; Pollak et al. 2003; Rauschecker et al. 1995; Wang et al. 1995) and birds (Gentner and Margoliash 2003; Grace et al. 2003; Langner et al. 1981; Leppelsack and Vogt 1976; Plummer and Striedter 2000; Scheich et al. 1977). In many instances, the emergence of higher-level selectivity can be understood in terms of combinations of lower-level selectivity found in presynaptic areas (i.e., in bats: Fitzpatrick et al. 1993; Xie et al. 2005; and in songbirds: Lewicki and Arthur 1996; Margoliash and Fortune 1992). In this study, we were interested in determining the role of the auditory thalamus in extracting information-bearing features of vocalizations and probing whether the neural code in the auditory thalamus is intermediate to the one found in the midbrain and pallium as supported by the hierarchical model, or potentially, more similar to the midbrain, as predicted by the classic "relay nucleus" theory of primary thalamic processing (see Sherman and Guillery 2002).

We studied these questions in the songbird model system, which offers unique opportunities. Songbirds are known for using complex and plastic vocalizations that are critical for their survival and reproductive fitness. In addition, the neural responses for natural sounds and vocalizations have been well characterized in the songbird auditory midbrain and primary and secondary auditory cortex-like regions. Moreover, they appear to be optimized for processing natural sounds (Hsu et al. 2004b; Woolley et al. 2005). However, the auditory thalamus has been conspicuously missing in past neurophysiological studies in songbirds in part because it has been difficult to record single unit activity from this small oval-shaped nucleus deep in the brain (some early work can be found in Bigalke-Kunz et al. 1987; Durand et al. 1992; Ströhmann et al. 1994). On the other hand, anatomical studies have shown that ovoidalis (Ov) is a central nucleus in the complex recurrent auditory network of the avian telencephalon (see Fig. 1 for a schematic based on Kelley and Nottebohm 1979; Vates et al. 1996; Zeng et al. 2004), and gene expression studies have implicated Ov in the processing of learned communication sounds (Brauth et al. 2007). In this study, we recorded neural activity in response to a large set of complex songs for a representative ensemble of Ov neurons in zebra finches and used the data to estimate the first spectrotemporal receptive fields (STRFs) from this nucleus. We were then able to determine whether the songbird auditory thalamus plays a distinct role in processing behaviorally relevant features of bird song by comparing the different types of STRFs found in Ov [analogous to mammalian medial geniculate body (MGB)] to those reported previously in the zebra finch's auditory midbrain, the mesencephalicus lateralis dorsalis [MLd: analogous to mammalian inferior colliculus (IC)], and the primary auditory pallium, field L (analogous to mammalian A1) (Nagel and Doupe 2008; Sen et al. 2001; Woolley et al. 2009).

Address for reprint requests and other correspondence: F. Theunissen, University of California, Berkeley, Psychology Department, 3210 Tolman Hall, Berkeley, CA 94720-1650 (E-mail: theunissen@berkeley.edu).

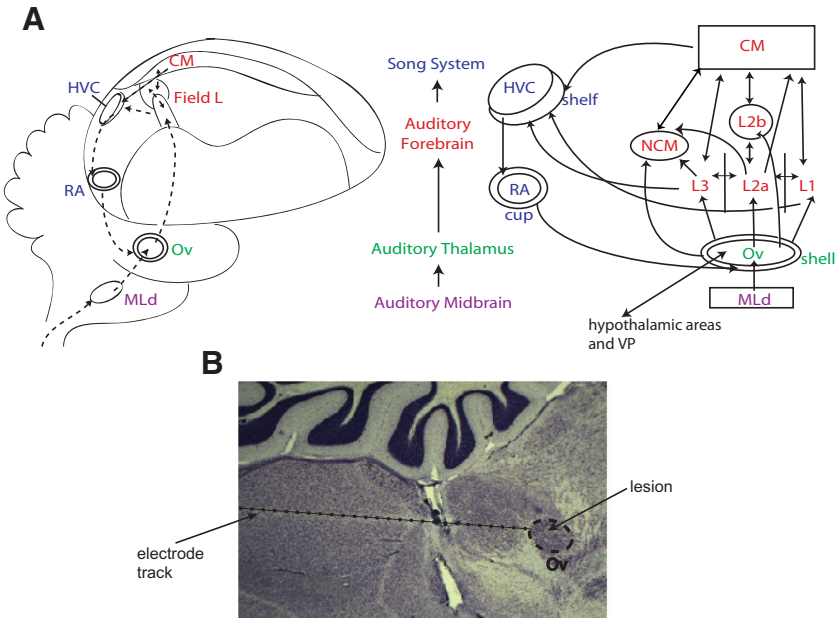

FIG. 1. *A*: schematized depiction of a sagittal section of the male zebra finch brain (to the left) and anatomical subdivisions and connectivity (to the right) of the avian auditory system in relation to the specialized song system with its shelf subregions. It is thought that the HVC-shelf/RA-cup subregions mediate the feedback signal from field L back to the ovoidalis shell. Note that not all feedforward connections pertaining to the song system are shown here. *B*: picture of a sagittal section of the male zebra finch brain (stained with cresyl violet) showing the electrode penetration, nucleus ovoidalis, and the lesion.

METHODS

Animal procedures

Eleven adult (>100 days old) male zebra finches (*Taenopygia guttata*) were used in all of our extracellular electrophysiological recording experiments. All subjects were raised by their parents in our zebra finch colony. Food and water were available to the birds at all times. Two days before the recording, we fixed a post to the bird. The subject was anesthetized with 0.03–0.05 ml Equithesin intramuscularly (0.85 g of chloral hydrate, 0.21 g of pentobarbital, 0.42 g of $MgSO_4$, 8.6 ml of propylene glycol, 2.2 ml of 100% ethanol, to a total volume of 20 ml with H_2O). The bird was immobilized in a stereotax with the aid of a beak holder (10°) and ear bars. After a local anesthetic of 2% lidocaine was administered on the scalp, a 1 mm square of the top layer of skull around the Ov reference (0.7 mm lateral from the midsagittal Y-sinus) was removed. A small stainless steel post was then fixed to the skull with dental cement.

On the day of the recording, the bird was anesthetized with three intramuscular injections of 25–30 μl each of 20% urethan administered at half hour intervals. We chose to perform our recordings under urethan anesthesia because we wanted to compare responses in the thalamus to our recordings obtained under similar anesthetic conditions in both the auditory midbrain and pallium (see Woolley et al. 2009). However, we recognize the possibility that response properties do change due to the anesthesia. In a study by Capsius and Leppelsack (1996) of field L responses of starlings both under anesthesia and in the awake bird, urethan reduces the spontaneous discharge rate significantly, decreases inhibition in all auditory subunits, and does not affect excitation much. More importantly, for field L subregions thought to receive direct inputs from the auditory thalamus, the "on" and sustained excitation changed only weakly. Based on this information, we hypothesize that the thalamic Ov responses might also be affected only weakly by urethan anesthesia, although this conjecture has to be tested. It is also probable that the anesthesia might get rid of some of the feedback connections thereby changing the receptive fields to a certain degree, and these differences would also have to be tested by recording in the awake bird and comparing them to the results obtained here.

To set up for the recording session, the bird's head was immobilized by attaching the steel post to a holder on the stereotax frame. A very small hole in the lower layer of the skull and the dura were made over the Ov reference. Extracellular tungsten electrodes (1–4 MΩ resistance) were lowered into the brain using a microdrive. The bird and the stereotax were then placed in a calibrated, double-walled anechoic chamber where a speaker was used to present the stimuli. The volume of the speaker was set to deliver zebra finch songs at 75 dB SPL peak (B&K Sound Level Meter, RMS weighting type B, slow weighting). The speaker was placed 20 cm in front of the bird's head. All animal procedures were approved by Animal Care and Use Committee at UC Berkeley.

Stimulus design and stimulus recordings

Conspecific song and white noise were used as search stimuli. To estimate the STRF, we used an ensemble of unfamiliar conspecific song (Con). This song stimulus ensemble consisted of 20 adult male zebra finch songs, each of which lasted ~2 s in duration and was recorded from a different bird in a sound-attenuated chamber. We have recently shown that 20 song exemplars allow for the diversity of spectral and temporal patterns occurring in zebra finch song such that the sampling error in both the power spectrum and the modulation spectrum is minimal (Singh and Theunissen 2003). All songs were filtered between 250 and 8,000 Hz and then normalized to have identical overall mean power during the nonsilent parts of the song.

Presentation of the different songs was randomized within a trial. Two seconds of background spontaneous activity was recorded prior to the presentation of each stimulus. A random interstimulus interval with a uniform distribution between 4 and 6 s was used.

Electrophysiology and experimental protocol

Neural recordings were conducted in a sound-attenuated chamber. Single, dual (defined as 2 units), and multiunit (defined as a small cluster of 3–5 neurons based on spike shapes) spike arrival times were obtained by thresholding the extracellular voltage trace with a window discriminator. Neural recordings were assessed to correspond to single units (vs. multiunits) by the following: possessing a high signal-to-noise ratio in the recordings (amplitude signal-to-noise ratio >5), monitoring the shape of the triggered action potentials on a digital oscilloscope with trace storage, and calculating the spike autocorrelation function post hoc. All spike autocorrelation functions from the visually determined single units showed the signature depression ~0 ms from postspiking inhibition: examination of the distribution of interspike intervals (ISI) showed that the average probability of

finding ISIs <1 ms for recording sites classified as single units was 0.5% while the expected value for Poisson neurons with same mean rates is 4.3%. The average song-driven firing rate of single unit recordings was 22.9 ± 2.4 spikes/s with an average background firing rate of 3.2 spikes/s. The recordings from the double and multiunit sites also had similar song-driven and background firing rates (double units: song-driven firing rate 20.8 ± 3.4 spikes/s, background firing rate of 2.1 spikes/s; multiunits: song-driven firing rate 20.8 ± 2.6 spikes/s, background firing rate of 3.8 spikes/s). z scores also were similar for single, dual, and multiunits [single units: z score mean = 4.35; double units: z score mean = 3.42; and multiunits: z score mean = 3.86; $F(112,2) = 1.02$; $P = 0.36$]. Given that there are no differences in mean rates and z-scores, and that the STRFs derived from single, dual, and multiunits were similar (see Table 1), we believe that most of the spikes in the dual and multiunits recordings came from a single dominant neuron. Thus we include all the data in this paper and will denote when the data are considered separately and when it is pooled.

At the end of the recording pass, two electrolytic lesions (100 μA for 5 s) 300 μm apart were made outside the auditory thalamus to allow for future recording site reconstruction.

Histology

At the end of the recording session, the bird was killed with an overdose of Equithesin and transcardially perfused with 0.9% saline, followed by 3.7% formalin in 0.025 M phosphate buffer. The skullcap was removed and the brain was postfixed in 30% sucrose and 3.7% formalin to prepare it for histological procedures. The brain was sliced parasagittally in 40 μm thick sections using a freezing microtome. Alternating brain sections were stained with both cresyl violet and silver stain, which were then used to visualize electrode tracks and electrolytic lesions (see Fig. 1*B* for a cresyl violet-stained section, showing the electrode track, nucleus, and lesion). Because the core and shell of nucleus Ov together comprise a small nucleus, it was difficult assigning with much certainty different recording sites to different parts of the nucleus. Thus we did not differentiate between anatomical subgroups when analyzing the data.

Data analysis

STRF CALCULATION. A normalized reverse correlation analysis was used to determine the relation between the stimulus and responses. This analysis yields the STRF, a model of a neuron's auditory tuning properties. The STRF calculation entails three steps. First, the log-intensity spectrogram of the sound stimulus (e.g., a sample of song) is cross-correlated with the average time-varying response to that stimulus obtained by averaging across the 10–14 trials to obtain the spike-triggered average. Second, the spike-triggered average is normalized by the autocorrelations of the stimulus. Third, a regularization-cross-validation procedure is used to effectively minimize the number of parameters that are fitted in the STRF estimation. Detailed descriptions of this STRF methodology are found in Theunissen et al. (2000, 2001), Woolley et al. (2005), and Gill et al. (2006).

TABLE 1. *Breakdown of the major Ov STRF types according to single, two, and multiunits*

	Single Unit	Two Units	Multiunits	Total N (% of Classified Neurons)
NB-T	8	2	−	10 (22.2)
NB-S	7	2	4	13 (28.8)
2B	2	1	−	3 (6.6)
FS	3	3	1	7 (15.5)
BB	5 + 1 2BB	2	3 + 1 2BB	10 + 2 (26.6)

Ov, ovoidalis; STRF, spectrotemporal receptive field; NB-T and NB-S, narrowband temporal and spectral, respectively; 2B, two band; FS, frequency sweep; BB, broadband. Note that percentages in the Total N column indicate percentages of classified neurons, thereby leaving out the unclassified neurons.

Once the STRF is obtained, it is validated on data that were not used in the STRF calculation. The similarity between the predicted response and the actual response (PSTH), measured using noise-corrected correlation coefficients (r), provides a measure of how well the STRF captures the tuning of a neuron. The correlation coefficient was calculated after smoothing the PSTHs with a hamming window of 7 ms. The value was corrected by considering the ceiling value that could be obtained for r given the neural variability measured across trials. Neurons for which the STRF gave predictions below $r = 0.2$ were excluded from the analysis. The strength and the reliability of the auditory response were also quantified independently of the STRF model by estimating the coherence between a single spike train (called R) and the actual time-varying mean response (called A). One can estimate this coherence (γ^2_{AR}) by calculating the coherence between two PSTHs, each PSTH being composed of half the trials

$$\gamma^2_{AR} = \left(1 - M \frac{1 - \sqrt{\left(\frac{1}{\gamma^2_{R_{1,M/2}\bar{R}_{2,M/2}}} \right)}}{2} \right)^{-1}$$

Here M is the number of trials, and $\gamma^2_{R_{1,M/2}\bar{R}_{2,M/2}}$ is the coherence between the two half-PSTHs. The coherence, γ^2_{AR}, measures the degree of "phase-locking" of the neural response to the sound as a function of frequency, ω. A single quantifier can be obtained by integrating over all frequencies and this number can be expressed in information units in bits/s

$$I = -\int_0^\infty \log_2[1 - \gamma^2(\omega)] d\omega$$

More detailed descriptions of the calculations of the normalized correlation coefficient and of the response coherence and information can be found in Hsu et al. (2004b).

EXTRACTING TUNING PARAMETERS FROM THE STRFs. As we did for the STRFs obtained in field L and MLd (Woolley et al. 2009), we extracted spectral and temporal tuning parameters and also the ratio of excitation to inhibition by fitting STRFs in Ov with a product of Gabor functions (Qiu et al. 2003)

$$\mathrm{STRF}(t, f) \approx \mathrm{AH}(t) \cdot G(f), \text{ where}$$

where

$$H(t) = e^{-0.5[(t - t_0)/\sigma_t]^2} \cdot \cos[2\pi \cdot \Omega_t(t - t_0) + P_t] \text{ and}$$

and

$$G(f) = e^{-0.5[(f - f_0)/\sigma_f]^2} \cdot \cos[2\pi \cdot \Omega_f(f - f_0) + P_f] \qquad (1)$$

Singular value decomposition (SVD) of matrices initially decomposes the STRF into separable time and frequency functions (an example is shown in Fig. 2). The fit of the parameters describing the Gabor function is performed by minimizing the mean square error between the H and G functions obtained from the SVD and those obtained from the Gabor model. The fitted parameters for the temporal function are t_0, the temporal latency (in seconds); σ_t, the temporal bandwidth (in seconds); Ω_t, the best temporal modulation frequency (in Hz); and P_t, the temporal phase. Similarly, the fitted parameters for the spectral function are f_0, the best frequency (in Hz); σ_f, the spectral bandwidth (in Hz); Ω_f, the best spectral modulation frequency (in cycles/Hz); and P_f, the spectral phase. Here we analyzed the spectral and temporal bandwidths, the ratio of excitation to inhibition [defined as (excitatory gain − inhibitory

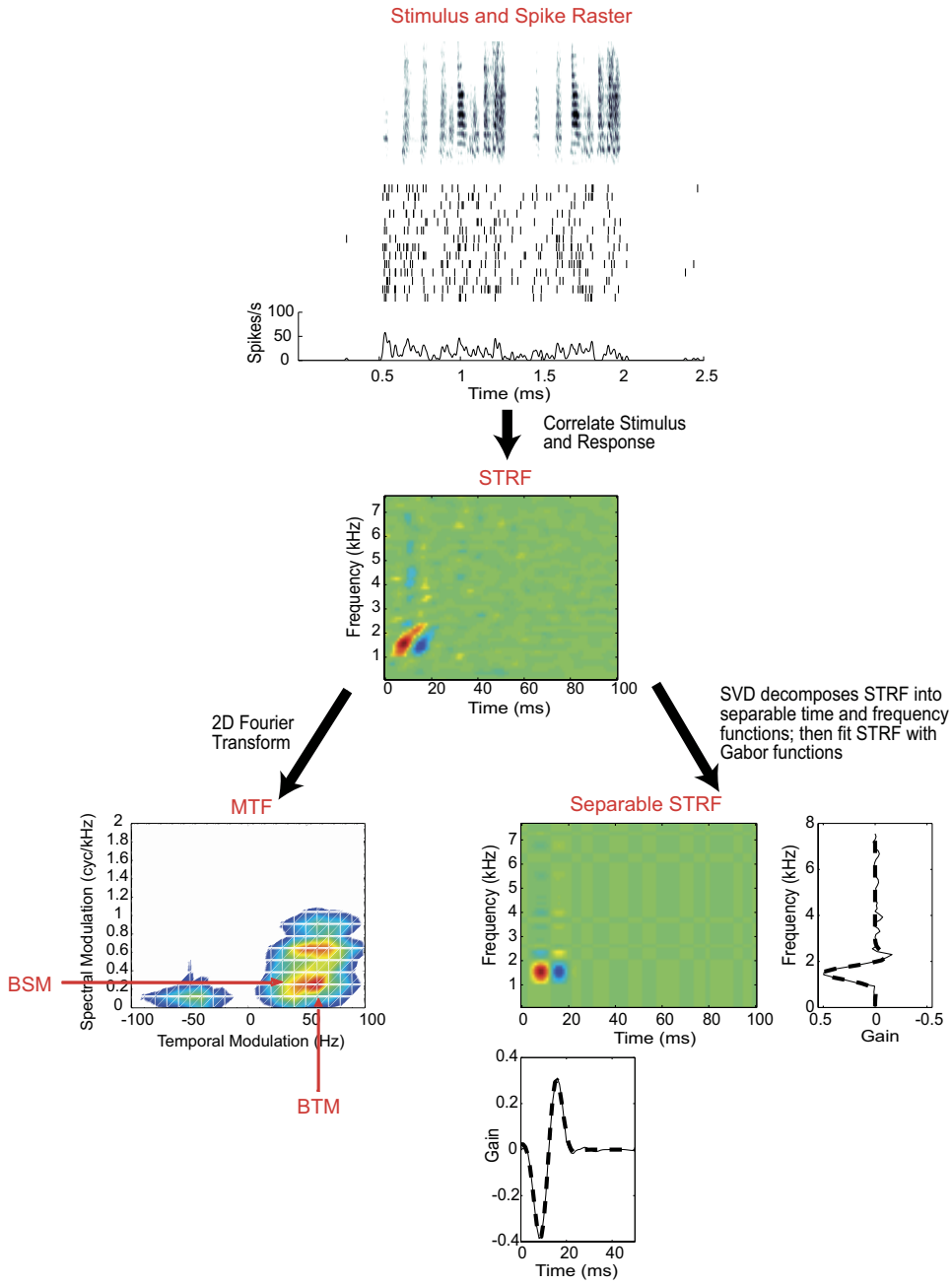

FIG. 2. *Top*: example of a conspecific song in spectrographic representation and the spike raster (14 trials) and peristimulus time histogram (PSTH) of a single ovoidalis (Ov) neuron in response to this stimulus. *Middle*: the spectrotemporal receptive field (STRF) for this particular cell is classified as a frequency sweep (FS). *Bottom, left*: the responses can also be plotted in modulation space. To calculate the modulation transfer function (MTF) of a STRF, a 2-dimensional (2D) Fourier fast transform (FFT) is performed on the STRF and then plotted on the 2D modulation Cartesian plot. Parameters like the best spectral modulation frequency (BSM) and best temporal modulation frequency (BTM) can be extracted from this modulation space. (Bottom, right): Singular value decomposition (SVD) decomposes the STRF into separable time and frequency functions. The fit of the parameters describing the Gabor function is performed, and parameters such as best frequency, latency and spectral and temporal bandwidths are extracted (see METHODS).

gain)/(excitatory gain + inhibitory gain)] and the latency. Note that in this analysis the latency corresponds to the time point of maximum response.

CLUSTER ANALYSIS. STRFs in MLd and field L can be categorized into functional groups based on their shape (Woolley et al. 2009). These functional groups are evident by visual inspection of the data but a more quantitative and objective grouping was performed by a clustering analysis using a genetic algorithm. In this study, we did not have enough data to perform a similar quantitative grouping based on clustering methods. Instead we used both a qualitative sorting, based on visual inspection, and a new set of quantitative approaches based on the similarity between neurons in Ov and neurons in MLd and field L. Together we were able to capture the nature and diversity of Ov neurons and also compare spectrotemporal tuning across the midbrain, thalamus, and pallium (our data here is compared with the data reported in Woolley et al. 2009).

MULTIDIMENSIONAL SCALING ANALYSIS. A nonclassical multidimensional scaling (MDS) analysis was performed on the similarities between STRFs from MLd, Ov, and field L (via built-in functions in Matlab). For the purposes of both raw pair-wise comparison of STRFs and also the MDS, the distance between two STRFs was taken to be 1 minus the correlation coefficient of the coefficients of the two STRFs after optimal alignment with arbitrary, wrap-around frequency shifts and with latency shifts $\leq$10 ms. We then performed statistics on both the original distances and the MDS distances. The objective of MDS is to represent data where we know the pairwise distances between points in a low-dimensional space. Generally, the original distances will have to be altered somewhat to fit in the lower-dimensional representation; MDS strives to minimize the distortion of the true inter-point distances while representing all points in the low-dimensional space. We projected neurons from MLd, Ov, and field L onto a two-dimensional (2D) manifold using MDS, respecting as much as possible the pairwise interneuron distances. The residual

stress of the MDS (i.e., the mean discrepancy between the true distances and those of the reduced MDS plot) was 17%, so a 2D plot gives a fairly good rough idea of how these neurons are grouped functionally.

In this MDS space, the variation of neurons was characterized in two ways: using the raw SDs of the points from MLd, Ov, and field L and using a Mahalanobis measure of the separation between STRFs from these brain areas. To obtain statistics about the difference between these areas, we used a bootstrap technique where we resampled (with replacement) 10,000 times from the data distribution.

ENSEMBLE MODULATION TRANSFER FUNCTION (EMTF) ANALYSIS. To supplement the MDS measure of the diversity of tuning, we also characterized the range and distribution of the peak gain in the modulation transfer functions (MTFs) of all neurons in MLd, Ov, and field L. Peak modulation tuning gains were determined by taking the amplitude of the 2D Fourier transform of the STRF to represent the STRF in the modulation tuning space (Kowalski et al. 1996; Qiu et al. 2003; Theunissen et al. 2004; Woolley et al. 2005). The temporal modulation and spectral modulation at the peak gain are also good estimates of the best temporal modulation (BTM: best temporal modulation) and the best spectral modulation (BSM: best spectral modulation) of the neuron. These values are very similar to the Ω_t and Ω_f, obtained from fitting the STRF with Gabor functions (see Fig. 2).

RESULTS

The goals of this study were one, to investigate whether the neural code in the auditory thalamus in songbirds can be interpreted in terms of a hierarchical model of sensory processing, or in terms of the more classic relay nucleus theory of primary thalamic processing; and two, to understand Ov's role in processing complex natural sounds. To do so, we will directly compare our data obtained in Ov to the data obtained from similar experiments in our laboratory from the upstream auditory midbrain nucleus MLd and downstream auditory pallial region field L (previously reported in Woolley et al. 2009).

Auditory responses in Ov

Given that in zebra finches Ov is a tiny nucleus deep in the brain, obtaining neural recordings were difficult. We were able to record from 59 recording sites in nucleus Ov from 11 birds. Five of these recording sites' responses and their corresponding STRFs are shown both in Figs. 2 and 3. As observed in the raster plots of Fig. 3, neurons in Ov generally have high background firing rates and high stimulus-driven firing rates and can have bursty or variable responses. Across the population of recorded single units, we found that Ov neurons had both higher firing rates and more variability in their auditory responses than in both MLd and field L. The average stimulus-driven firing rates (SR) for single units were: SR = 9.1 ± 6.3 spikes/s for MLd, SR = 23 ± 16 spikes/s for Ov; and SR = 10.7 ± 6.5 spikes/s for field L; [$F(2,301) = 39.06$; $P < 10^{-4}$]. The background rate (BR) was also significantly higher in Ov [MLd: BR = 0.15 ± 0.46 spikes/s; Ov: BR = 3.25 ± 4.05 spikes/s; field L: BR = 0.62 ± 1.39 spikes/s; $F(2,301) = 46.71$, $P < 10^{-4}$]. We also assessed the reliability of the auditory evoked responses (also loosely referred to as phase-locking) by estimating the coherence between single spike trains and an estimate of the average time-varying mean firing rate. The average coherence can then be expressed as an information number in bits/s (see Hsu et al. 2004b). Neurons with a time-varying neural response that is significantly modulated by sound features and that show a high degree of spike timing reliability will have high information values. In comparison with MLd and field L, neurons in Ov had the lowest information values (Info): Info = 45.4 ± 46.2 bits/s for MLd;

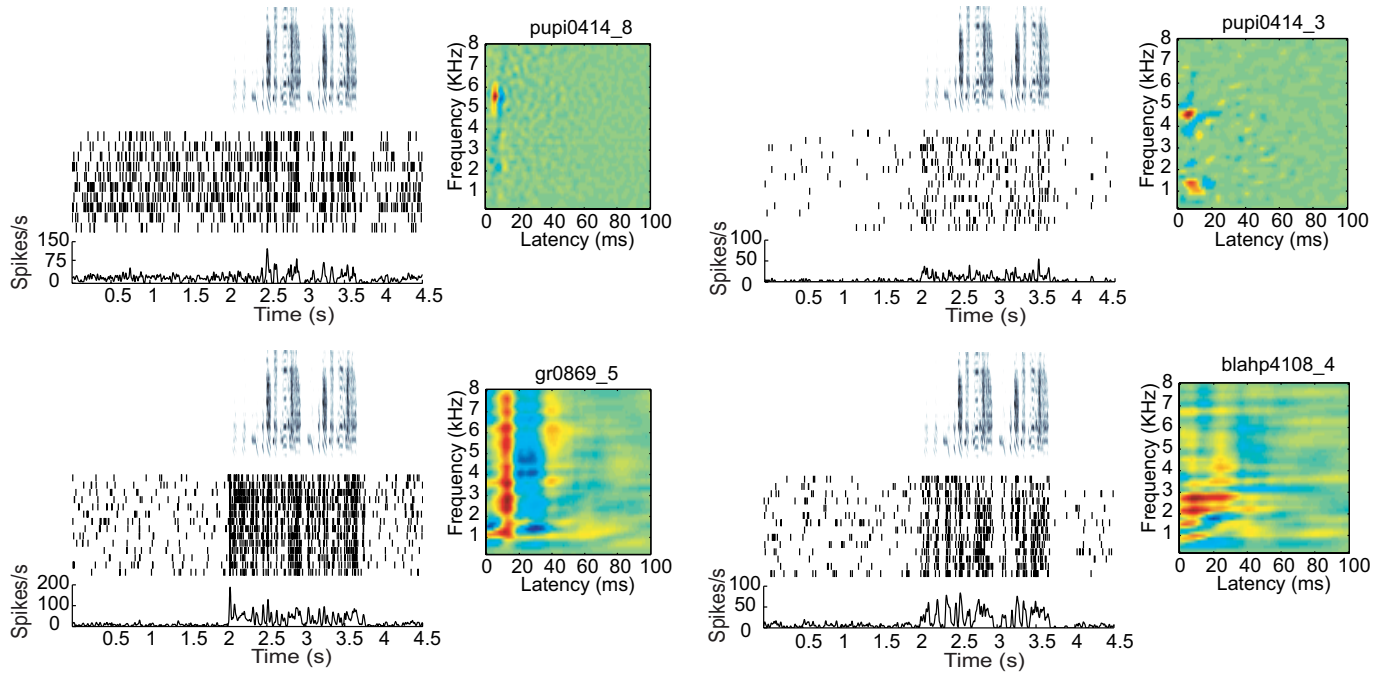

FIG. 3. Four distinct exemplars of single units in Ov from 4 different birds responding to the same conspecific song (as denoted by the song spectrogram on the *top panel,* spike rasters for 10–14 trials in response to the song, and corresponding PSTHs) and their corresponding STRFs. We chose to show 2 STRF types that were common in Ov and discussed in this paper (NB-T on *top left,* and BB on *bottom left*) and 2 STRF types that were not as abundant (2B on *top right*) nor typical (complex/unclassified on *bottom right*).

Info = 27.2 ± 44.3 bits/s for Ov; and Info = 32.0 ± 34.2 bits/s for field L [$F(2,326) = 5.30$, $P = 0.005$]. Therefore both in terms of firing rates and variability, neurons in Ov cannot be considered as having intermediate response properties between those in MLd and field L.

STRF types in the auditory thalamus

We estimated the STRF from the responses to 20 different conspecific songs for all 59 recording sites in our data set. Of these 59 STRFs, 54 gave predictions above $r = 0.2$ and therefore were included in our analysis. We then visually classified the STRFs based on their specific spectrotemporal patterns and using the nomenclature that we used previously to describe auditory tuning and function in MLd and field L (see Woolley et al. 2009). Of the 54 neurons, 9 neurons had STRFs that were too complex and idiosyncratic to be assigned to one particular functional class (see Fig. 3 for an exemplar complex or unclassified STRF). The remaining 45 neurons belonged to three major functional groups (see Fig. 4): narrowband (NB), broadband (BB), and frequency sweep neurons (FS).

As in field L and MLd (see Woolley et al. 2009), the STRFs of the NB neurons in Ov were characterized by very narrow frequency tuning and generally fast temporal tuning. The NB neurons could be further subcategorized into narrowband-temporal neurons (NB-T: 10/45) and narrowband-spectral neurons (NB-S: 13/45). The NB-T neurons were characterized by a rapid succession along the temporal dimension of positive and negative gain making them sensitive to sound onsets at a particular frequency. This temporally sharp subgroup of narrowband neurons were also commonly found in both MLd and field L (Woolley et al. 2009). Mean best frequencies (BF) estimated from the STRF (f_0 in *Eq. 1*) for NB-T type neurons were similar in MLd and Ov but were significantly different between field L and the lower areas of Ov and MLd [BF = 4.01 ± 1.43 kHz for MLd; BF = 4.07 ± 1.06 kHz for Ov; BF = 2.01 ± 0.53 kHz for field L; $t(32) = -0.104$, $P = 0.917$ for MLd and Ov comparison; $t(17) = 5.431$, $P < 0.001$ for Ov and field L comparison; $t(33) = 4.287$, $P < 0.001$ for MLd and field L comparison]. NB-T neurons in all three areas had similar mean spectral and temporal bandwidths (σ_f and σ_t in *Eq. 1*): $\sigma_f = 0.39 +- 0.24$ kHz in MLd; $\sigma_f = 0.37 \pm 0.19$ kHz in Ov; and $\sigma_f = 0.22 \pm 0.07$ kHz in field L; [$F(2,41) = 1.30$, $P = 0.28$); $\sigma_t = 3.7 \pm 1.4$ ms in MLd; $\sigma_t = 3.3 +- 0.9$ ms in Ov; and $\sigma_t = 4.6 \pm 2$ ms in field L; ($F(2,41) = 2.25$, $P = 0.11$]. The NB-S neurons in Ov were similar to NB-T neurons with the additional property that their STRFs revealed inhibitory sidebands in the spectral dimension. NB-S neurons were not found in MLd but were observed in both Ov and field L. Similar best frequencies were found for both Ov and field L NB-S neurons [BF = 2.22 ± 1.05 kHz for Ov; BF = 1.88 ± 0.37 kHz for field L; $t(19) = 0.889$, $P = 0.384$]. Ov and field L NB-S neurons also had similar spectral bandwidths ($\sigma_f = 0.36 \pm 0.16$ kHz for Ov; $\sigma_f = 0.23 \pm 0.09$ kHz for field L; $t(19) = 1.96$, $P = 0.06$) but the NB-S neurons in field L had longer temporal bandwidths [$\sigma_t = 3.9 \pm 0.9$ ms for Ov; $\sigma_t = 9.9 \pm 6.4$ ms for field L; $t(19) = -3.36$, $P = 0.003$]. Moreover, ~80% of NB-S neurons in Ov also had inhibitory gain temporally following the area of maximum excitatory gain (as NB-T neurons), whereas none of the NB-S neurons in field L showed this temporal inhibition. We also found three neurons in Ov that had two spectrally separated narrow excitatory fields in both time and frequency bands, and we called them 2Band neurons (2B: 3/45). Similar STRFs had been observed for a small subset of neurons in MLd (see Table 2) but with much wider frequency bands. Even though we did not record from any 2B neurons in field L, others have reported them in other forebrain areas and also nucleus Ov of barn owls (Perez

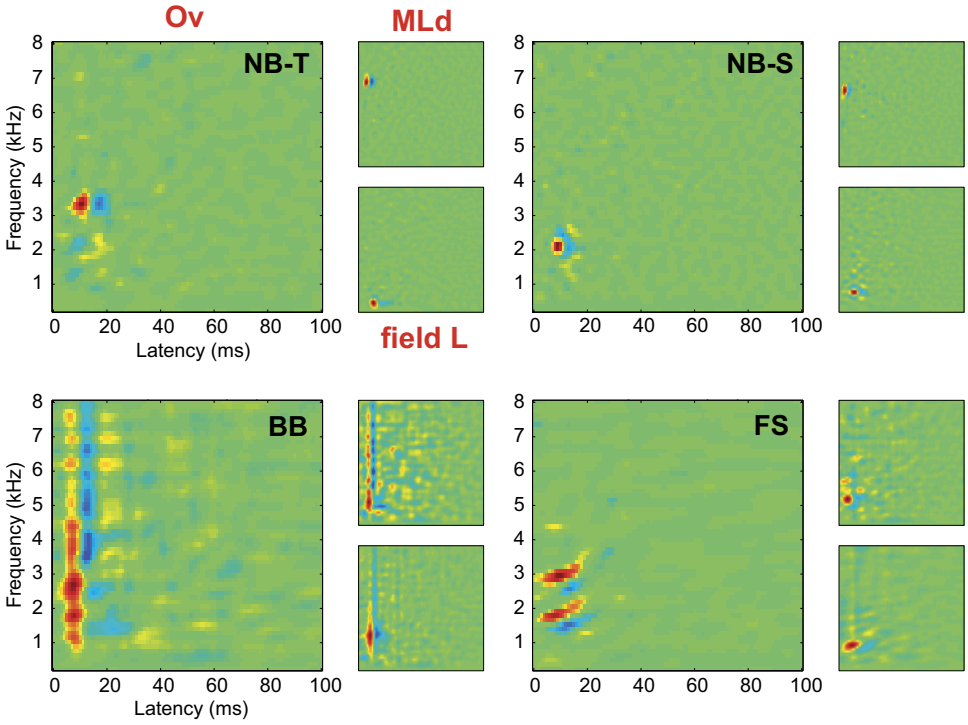

FIG. 4. Typical STRF types found in Ov (larger *inset* on left) and their closest matches found in MLd and field L (smaller insets on the right) as calculated by a correlation coefficient. The main functional types in Ov include: NB-T, narrowband, temporal; NB-S, narrowband-spectral; BB, spectrally broad band; and FS, frequency sweeps. Note that for the NB-S case, the closest match of a field L cell to an Ov cell was a regular NB cell without the longer spectral tail since field L NB-S cells lack the temporal inhibition lag.

TABLE 2. *STRF types in MLd, Ov, and field L*

	MLd (Clustered = 110/137)	Ov (Clustered = 45/54)	Field L (Clustered = 105/137)
NB-T, %	22.73	18.52	9.52
BF, kHz	4.01 ± 1.43	4.07 ± 1.06	2.01 ± 0.53
Spectral BW, kHz	0.39 ± 0.24	0.37 ± 0.19	0.26 ± 0.08
Temporal BW, ms	3.73 ± 1.41	3.34 ± 0.90	2.81 ± 0.54
Excit-inhib ratio	0.11 ± 0.08	0.05 ± 0.03	0.06 ± 0.06
NB-S, %	—	25.93%	7.62%
BF, kHz		2.22 ± 1.05	1.88 ± 0.37
Spectral BW, kHz	—	0.36 ± 0.16	0.23 ± 0.09
Temporal BW, ms		3.92 ± 0.89	9.90 ± 6.40
Excit-inhib ratio		0.04 ± 0.03	0.15 ± 0.14
2B, %	4.55	7.41	—
BF, kHz	1.55 ± 0.19	2.54 ± 1.48	
Spectral BW, kHz	1.65 ± 2.80	1.65 ± 2.54	
Temporal BW, ms	5.23 ± 4.60	11.37 ± 11.19	
Excit-inhib ratio	0.34 ± 0.18	0.23 ± 0.30	
Off, %	—	—	6.67%
BF, kHz			1.39 ± 0.40
Spectral BW, kHz			1.61 ± 0.78
Temporal BW, ms			8.76 ± 3.18
Excit-inhib ratio			0.07 ± 0.11
FS, %	—	14.8	2.91 (subset of hybrid cluster)
BB, %	67.27	24.07	39.05
BF, kHz	3.57 ± 1.01	2.52 ± 0.81	3.15 ± 1.36
Spectral BW, kHz	2.22 ± 0.48	1.71 ± 1.48	2.18 ± 0.64
Temporal BW, ms	5.35 ± 3.29	4.78 ± 1.88	5.78 ± 3.71
Excit-inhib ratio	0.10 ± 0.16	0.27 ± 0.33	0.02 ± 0.08
WB, %	5.45	—	11.43
BF, kHz	1.58 ± 0.10		2.73 ± 1.48
Spectral BW, kHz	0.74 ± 0.28		2.67 ± 4.04
Temporal BW, ms	15.02 ± 4.36		18.81 ± 5.36
Excit-inhib ratio	0.43 ± 0.22		0.32 ± 0.25

MLD and field L data is from Woolley et al. (2009). The percentages indicate percentages of classified neurons. Also slow narrowband cells are not reported here in the table, but see Woolley et al. (2009) for more details on this subgroup of narrowband cells (~16.19% of clustered STRFs in field L). Bandwidths (BW) not reported for FS cells because simple spectral and temporal bandwidths have little meaning for these cells. MLd, mesencephalious lateralis dorsalis; WB, wideband; off, offset; BF, best frequency; excit-inhib, excitatory-inhibitory.

and Pena 2006; Proctor and Konishi 1997; Vonderschen and Wagner 2009).

The second main group consisted of the broadband neurons (BB: 10/45), defined by their broad spectral tuning. Here, unlike in the case of NB-T neurons, the BFs for only the MLd and Ov comparisons were significantly different [BF = 3.57 ± 1.01 kHz for MLd; BF = 2.52 ± 0.81 kHz for Ov; BF = 3.15 ± 1.36 kHz for field L; $t(84) = 3.426$, $P < 0.001$ for MLd and Ov comparison; $t(51) = -1.519$, $P = 0.134$ for Ov and field L comparison; $t(113) = 1.904$, $P = 0.059$ for MLd and field L comparison]. As in MLd and field L, the spectral bandwidth of this group of neurons in Ov belonged to a separate mode: all these neurons had broad spectral tuning with relatively little variance and higher mean than those found for NB neurons. The spectral bandwidth was similar in all three areas [$\sigma_f = 2.2 \pm 0.5$ kHz for MLd; $\sigma_f = 1.85 \pm 1.6$ kHz for Ov; $\sigma_f = 2.18 \pm 0.6$ kHz for field L; $F(2,124) = 2.92$; $P = 0.06$]. As in MLd and field L, the temporal bandwidth of BB neurons in Ov exhibited a large amount of spread. In other words, BB neurons could be fast or slow. The mean temporal bandwidth as well as the range of temporal bandwidths within the BB neurons were similar in all three regions [$\sigma_t = 5.3 \pm 3.3$ ms for MLd; $\sigma_t = 5.7 \pm 4.1$ ms for Ov; and $\sigma_t = 5.8 \pm 3.7$ ms for field L; $F(2,124) = 15.17$, $P = 0.30$]. We also found two neurons that we called double broadband (2BB) as they were characterized by a broad spectral excitatory tuning followed by another equally similar broad spectral tuning. We did not find such neurons in MLd or field L, but these neurons in Ov are rare as well.

The third main group of STRFs was characterized by excitatory tuning increases in frequency over time and thereby categorized as frequency sweeps (FS: 7/45). FS STRFs had a mean BF of 2.14 ± 0.59 kHz. Although neurons with some frequency modulation were also observed in field L (and often classified as part of the hybrid class or HY), their frequency modulation was not as pronounced (as quantified in the following text with MTFs). Finally, 9/54 neurons' tuning was complex, a few of which consisted of multiple smeared excitatory frequency bands and a couple that were interspersed by inhibitory bands, followed by more excitatory bands. They were thereby unclassifiable into these three main categories. In Fig. 4, we show example neurons for these major types and their closest match (as determined by their similarity index) in MLd and field L.

We also wanted to compare the effectiveness of STRFs across the three nuclei: we estimated how much of the response of the neurons can be explained with the linear STRF model by calculating the correlation coefficient between the predicted response and actual response spike rate for all three nuclei. The quality of the predictions was roughly in the same range across these brain regions but the

mean differences were significant. The population of recorded Ov neurons was more linear than that recorded in MLd, which in turn was more linear than the population of recorded neurons in field L [MLd: $r = 0.56 \pm 0.14$; Ov: $r = 0.65 +- 0.14$; field L: $r = 0.46 \pm 0.15$; $F(2,326) = 34.07$, $P < 0.001$]. One might expect the performance of the linear model to decrease in higher sensory areas with the emergence of more complex responses. This trend is true for field L relative to MLd or Ov, but we also found higher linearity in Ov than in MLd. This discrepancy can be explained in part by the higher firing rates of Ov neurons, as compared with MLd neurons, and the fact that neurons with higher firing rates (and higher spontaneous rates) generally have more linear stimulus-response functions.

In summary, NB-T and BB neurons found in Ov were similar to those found in MLd and field L. The NB-S neurons, while absent in MLd, are present in Ov and show some similarity with the NB-S neurons in field L but are also distinct in that the Ov ones are faster and detect temporal changes. NB-S neurons in Ov could therefore be an intermediate step from NB-T neurons in MLd and NB-S neurons in field L. On the other hand, the FS neurons found in Ov could not be considered as simple linear intermediates between neurons found in MLd and field L because they show more pronounced frequency sweep tuning. We discuss this further in the DISCUSSION.

Table 2 summarizes the different STRF types across the three auditory regions, revealing that the auditory thalamus's receptive fields show both overlap with STRFs found in the midbrain and the pallium as well as distinctive features that can be considered to be either intermediate or unique to Ov. We also report the spectral bandwidths, temporal bandwidths, and the ratio of excitatory-inhibitory bands for each class. The similarity between MLd, Ov and field L will be further quantified in the next section and the potential relevance of these STRF types for encoding vocalizations will be discussed in greater detail in the DISCUSSION.

Is the auditory thalamus more similar to the auditory midbrain or pallium?

We calculated pair-wise distances between all Ov and MLd STRFs and similarly between all Ov and field L STRFs (see METHODS). For each Ov STRF, we then chose the closest MLd STRF and field L STRF. We then performed a paired *t*-test on these closest pair-wise differences and found that Ov STRFs matched more closely to field L STRFs than to MLd STRFs: mean difference of Ov-MLd STRFs = 0.35; median difference of Ov-MLd STRFs = 0.34; mean difference of Ov-field L STRFs = 0.33; median difference of Ov-field L STRFs = 0.32; $t(50) = 2.48$, $P = 0.01$. To help visualize these distances, we plot the best pair-wise similarity index (SI) for the Ov-MLd STRF comparison versus the Ov-field L STRF comparison (Fig. 5). As expected from the results of the *t*-test, a greater number of points lie below the unity line denoting that a greater number of Ov STRFs are similar to those found in field L than in MLd. On the same figure, we also show typical examples of STRFs of: neurons that are unique to Ov and therefore have a low SI to neurons in MLd and L (left margin); neurons that are also found in MLd and field L and therefore have high SI's for both comparisons (right margin); and neurons found in both Ov and field L but not in MLd and therefore have a high SI for the field L comparison and a low SI for the MLd comparison (bottom margin). The example STRF that is unique to Ov shows sensitivity for frequency sweeps that is best matched (but poorly) by the simple NB-T neurons found in MLd and field L. The example STRF that can be found in all three regions is the NB-T neuron (simple BB neurons are also found in all 3 regions). The example STRF found in Ov that is similar to field L's hybrid STRF (characterized by a combination of excitatory NB tuning in low frequencies, followed by inhibitory BB tuning) is the FS cell (but with more pronounced frequency modulation), and neither of these are found in MLd. NB-S neurons were also unique to field L and Ov.

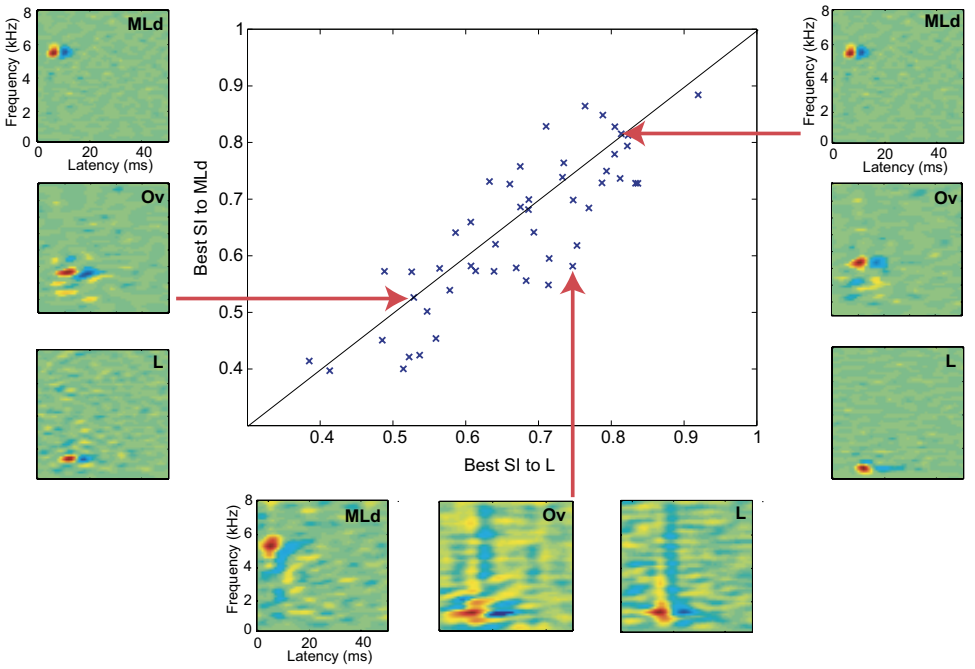

FIG. 5. A scatterplot of best pairwise similarities between each Ov STRF and its counterpart MLd STRFs (on *y* axis) and each Ov STRF and its counterpart field L STRFs (on *x* axis). A greater number of points lie below the unity line denoting that a greater number of Ov STRFs are similar to those found in field L than to those found in MLd. Three sets of exemplar STRFs from each of the regions are shown here: Ov neurons that are more unique in tuning as compared with MLd or field L (as illustrated by a FS STRF from Ov compared with its closest NB-T STRFs from both MLd and field L, on the left); Ov neurons that are more similar to field L than to MLd (as illustrated by a FS STRF from Ov compared with its closest NB-T STRF from MLd and a HY STRF from field L, on the bottom – note that the Ov FS cell is "sweepier" than the HY cell); and Ov neurons that are similar to both MLd and field L (as illustrated by similar NB-T STRFs across the 3 nuclei, on the right).

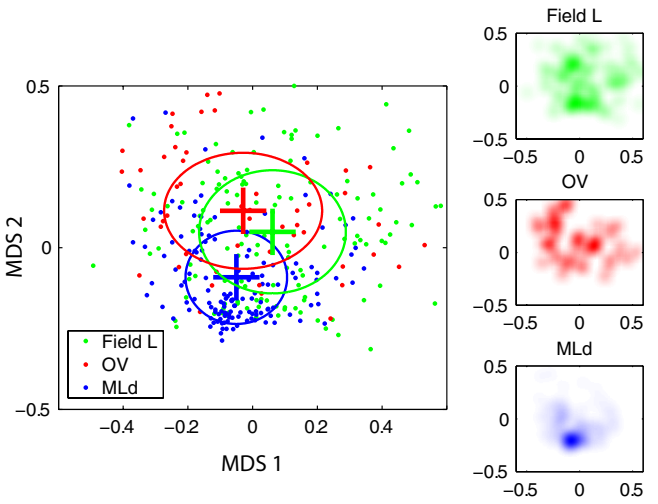

FIG. 6. Projections of neurons from MLd, Ov, and field L onto a 2-dimensional manifold using multidimensional scaling (MDS), respecting as much as possible the pairwise inter-neuron distances (see METHODS). The ovals correspond to 1 SD. The subpanels to the right depict Gaussian-weighted points to help visualize the density of data points in each area.

We also performed a MDS analysis on the distances between the STRFs of all neurons from MLd, Ov, and field L together. The MDS takes a matrix of pairwise differences and plots each point in a hyperplane in such a way that it minimizes "stress," that is the discrepancy between what the distances actually are and how they appear in the hyperplane (see METHODS). In this analysis, we found that two dimensions were sufficient to capture a large amount of the similarity as quantified by the distance between STRFs (the residual stress was 17%). As seen in Fig. 6, the neurons in Ov are more diverse in their functional properties than neurons in MLd. This functional heterogeneity is evident in the spread of the points in the scatter pot (also shown by the ovals that measure 1 SD). On the other hand, Ov and field L appear to have similar spread (also shown by Gaussian-weighted points in the subpanels on the right to help visualize the density of data points from each area). This result is statistically significant: MLd has a lower variance in the MDS representation than either L ($P < 0.0002$; 2-tailed bootstrap with 10,000 resamplings) or Ov ($P < 0.0002$); however, field L and Ov have approximately the same variance ($P = 0.6$), which is 97% greater than that in MLd.

Additionally, one can also see from Fig. 6 that the center of the cloud of points corresponding to Ov neurons is closer to the center of points corresponding to field L neurons than to the center corresponding to MLd neurons. We can quantify this effect by calculating the distance between the centers of field L, Ov, and MLd cells in the MDS representation using either the Mahalanobis distance or the Euclidian distance in the MDS space. With the Euclidian metric, Ov and field L are 45% closer than Ov and MLd ($P = 0.007$, determined by bootstrap); using the Mahalanobis distance, the same discrepancy is 62% ($P = 0.0002$). In absolute terms, the Mahalanobis distance between Ov and field L is 0.51, between field L and MLd is 1.02 and between Ov and MLd is 1.40. The Mahalanobis distance between Ov and MLd is greater than that between field L and MLd primarily because the MDS direction along which Ov maximally varies (horizontal in Fig. 6) is nearly orthogonal to the direction separating MLd and Ov (vertical in Fig. 6). This analysis shows that Ov neurons, taken as an ensemble, are distinct from both MLd and field L but more similar to field L than to MLd neurons.

We also quantified differences in peak MTFs for all MLd, Ov, and field L units. The peak MTFs for all the data are shown in Fig. 7. Peak MTFs of STRFs from MLd tend to congregate on the temporal frequency axis between 40 and 100 Hz (also seen in the MLd subpanel of Fig. 7), a result of the high number of BB neurons with fast temporal properties found in the auditory midbrain. In contrast, neurons in Ov and field L show more tuning diversity and include many more neurons with slower temporal properties but higher spectral modulations corresponding to sharper spectral tuning (and this is quantified above). On average, peak best temporal modulation frequencies (BTM) were significantly higher in MLd than in both Ov and field L [BTM = 63.62 ± 25.17 Hz for MLd; BTM = 50.72 ± 30.94 Hz for Ov; BTM = 43.02 ± 27.49 Hz for field L; $t(187) = 2.91$, $P = 0.003$ for MLd and Ov comparison; $t(190) = 1.64$, $P = 0.10$ for Ov and field L comparison; and $t(279) = 6.54$, $P < 0.0001$ for MLd and field L comparison]. Conversely, peak best spectral modulations (BSM) were significantly lower in MLd than in both Ov and field L [BSM = 0.10 ± 0.22 1/kHz for MLd; BSM = 0.36 ± 0.44 1/kHz for Ov; BSM = 0.31 ± 0.45 1/kHz for field L; $t(187) = -5.18$, $P < 0.0001$ for MLd and Ov comparison; $t(190) = 0.66$, $P = 0.50$ for Ov and field L comparison; and $t(279) = -4.77$, $P < 0.0001$ for MLd and field L comparison].

Figure 7 also shows that there is an abundance of neurons in Ov that are tuned for intermediate temporal modulations (20–60 Hz) and intermediate spectral modulations (0.2 to 1.2 cycles/kHz; see the Ov subpanel in Fig. 7 for a clearer visual depiction). These points correspond to neurons selective for frequency sweeps. Compared with field L, neurons in Ov tuned to spectral modulations were more common at intermediate temporal modulations compared with low temporal modulations, while the opposite is true for field L neurons. We quantified this effect by analyzing the values of the temporal

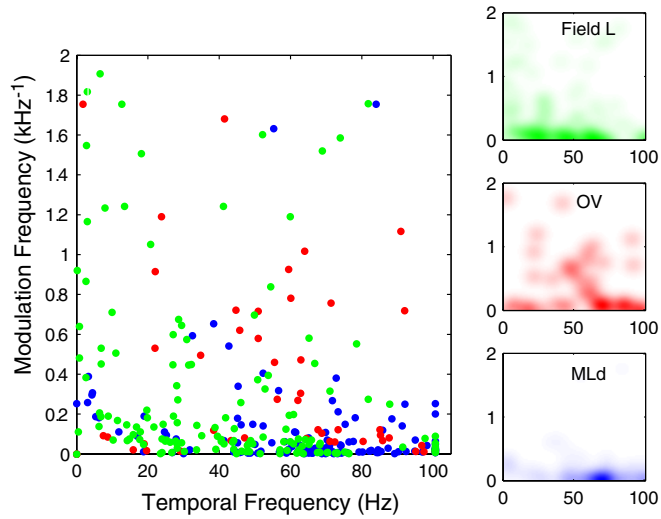

FIG. 7. Distribution of the peaks of the modulation transfer functions of neurons in MLd, Ov, and field L (see METHODS). Each dot shows the peak of the modulation transfer function of 1 STRF in terms of BTM and BSM. Only 1 quadrant (positive spectral and temporal modulations) of data points is shown here. The subpanels to the right depict Gaussian-weighted points to help visualize the density of data points in each area.

modulations for neurons that had spectral modulations > 0.2 cycles/kHz. For these neurons, we found an overall difference in mean temporal modulations between Ov and field L [BTM at cutoff = 49.14 ± 34.33 Hz for MLd; BTM at cutoff = 51.22 ± 22.05 Hz for Ov; BTM at cutoff = 33.10 ± 26.89 Hz for field L; $t(38) = -0.22$, $P = 0.81$ for MLd and Ov comparison; $t(66) = 2.70$, $P = 0.008$ for Ov and field L comparison; $t(64) = 2.02$, $P = 0.04$ for MLd and field L comparison]. The cutoff at 0.2 cycles/kHz (corresponding to spectral modulations with periods <5 kHz) was chosen to exclude BB neurons from this analysis (see Fig. 4 from Woolley et al. 2009). BB neurons can be tuned to fast temporal modulations, but given their large spectral bandwidth, they are not tuned for frequency sweeps.

These results, as quantified by raw pair-wise differences, MDS, and peak modulation transfer functions, are consistent with the notions that: MLd has less diversity in receptive field properties than Ov and field L; Ov and field L are more similar to each other than Ov and MLd; and neurons in Ov might show some degree of specialization that is absent in MLd and also distinct from that of field L.

Latencies increase from midbrain to thalamus to pallium

As expected, the mean latencies increase by a few milliseconds from MLd (7.1 ms, see Fig. 8*A*) to Ov (10.2 ms) to field L (13.9 ms); mean latencies are significantly different ($P < 0.0001$, 2-sided Wilcoxon rank sum test). The latencies of thalamorecipient subfield L2a cells (mean: 12.4 ms) are slightly shorter than those of nonL2a cells (14.9 ms), and this difference is also significant ($P = 0.05$, rank sum test). The mean latencies across these three nuclei along the hierarchy substantiate the expected feedforward connection, but the tails of the latency distributions (as seen in Fig. 8*A*) also point to a

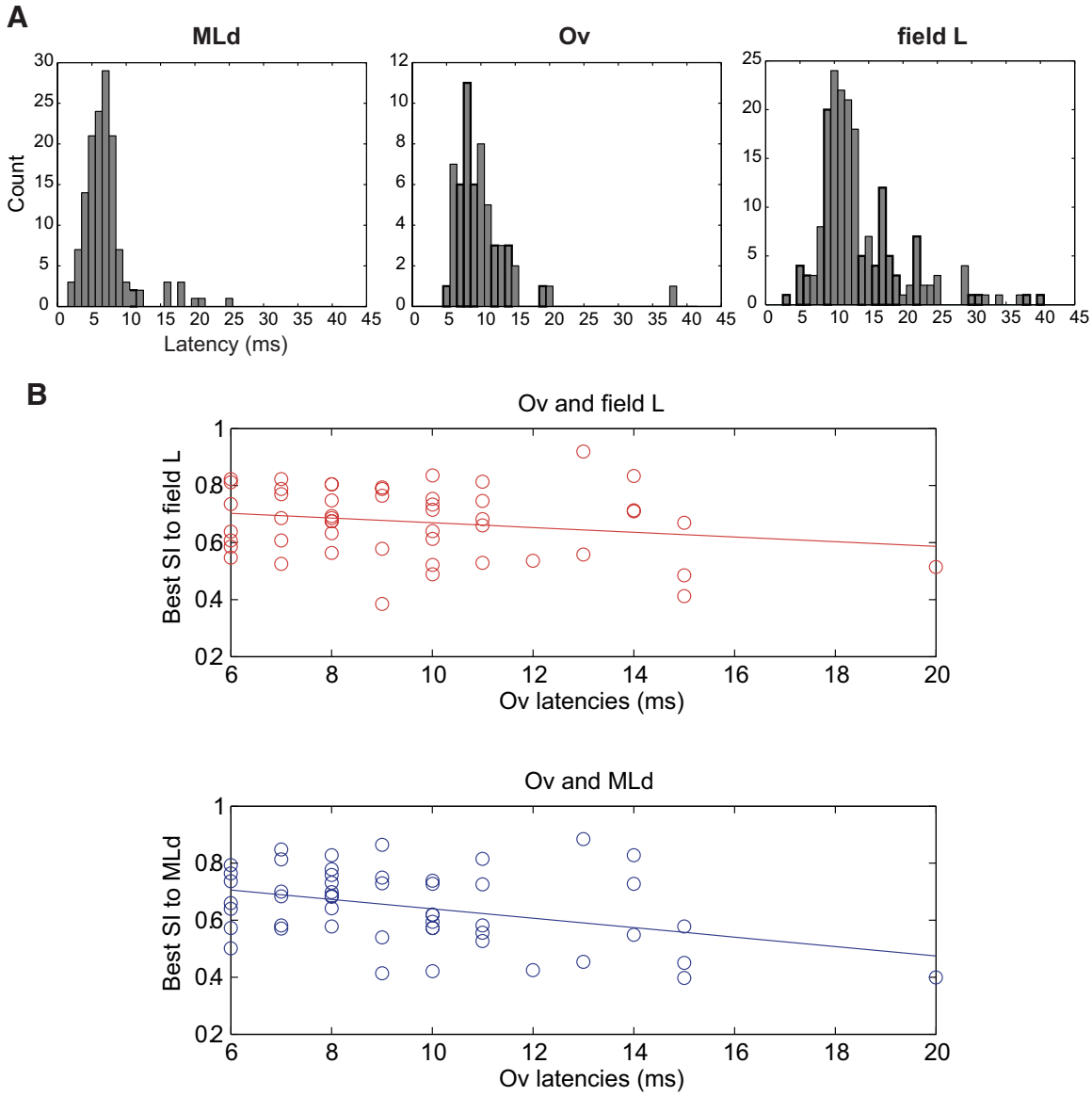

FIG. 8. *A*: Distributions of latencies (in ms) derived from STRFs for MLd, Ov, and field L. The spread of these distributions suggests both feed forward and feed back projections are possible, since the fastest field L neurons fire before the slowest MLd neurons. *B*: plot of best SI of Ov STRFs to field L STRFs vs. latency (in ms) derived from Ov STRFs (*top*), and plot of best SI of Ov STRFs to MLd STRFs vs. latency (in ms) derived from Ov STRFs (*bottom*). Earlier Ov responses are more similar to MLd STRFs, whereas latter Ov responses are more similar to field L STRFs.

possible role of feedback at each level because a substantial portion of field L neurons have faster latencies than some neurons in both MLd and Ov. If feedback from field L is indeed playing a role in shaping Ov responses, then we would expect the earlier Ov responses to be more similar to MLd (due to faster feedforward connections), whereas the latter responses might be more similar to field L STRFs due to possible feedback from field L. To test this prediction, we plotted the best SI of Ov STRFs to MLd STRFs versus latency derived from Ov STRFs, and similarly we plotted the best SI of Ov STRFs to field L STRFs versus latency derived from Ov STRFs (Fig. 8*B*). A correlation analysis suggests that, as predicted, the STRFs of Ov neurons with shorter latencies are more similar to MLd STRFs than the STRFs of Ov neurons with longer latencies ($r^2 = 0.142$, $F = 7.944$, $P = 0.006$). On the other hand, that same relationship (or more precisely to the same significant degree) is not found when Ov neurons are compared with field L neurons (where the correlation is not significant, $r^2 = 0.042$, $F = 2.149$, $P = 0.149$). This is an important control because it shows that the effect is not simply due to the fact that longer latency neurons have more complex neurons that are more idiosyncratic and therefore less similar to the STRF of any other neuron.

DISCUSSION

The avian auditory thalamic nucleus, Ov, is situated downstream of the auditory midbrain nucleus, MLd, and upstream of the auditory pallial region, field L, in the ascending auditory processing stream. Some of the response properties of Ov neurons such as latency reflect this intermediate position. The analysis of Ov auditory receptive fields also showed a fraction of neurons have STRFs either preserved from those found in MLd or intermediate between those found in MLd and field L. However, Ov also exhibits unique response properties as well as a diverse set of complex STRFs. For instance, Ov neurons have higher firing rates and higher neuronal variability than do either MLd or field L neurons. The range and complexity of Ov STRFs are also more similar to those found in the pallium than in the midbrain, and they also exhibit some unique features, in that Ov has a higher proportion of neurons with sensitivity for frequency sweeps.

Different STRF types in Ov encode for distinct features of vocalizations

The STRF, as the best linear model that transforms any time-varying stimulus into a prediction of the firing rate of a neuron, can help explain some of the encoding properties of complex auditory neurons for a subset of specialized sounds, such as vocalizations. Because we wanted to understand both the role of the auditory thalamus in processing complex sounds and its role in the hierarchical auditory processing stream, we visually classified Ov neurons according to their joint spectrotemporal properties and used similar functional divisions and nomenclature as we did for neurons in MLd and field L, as in Woolley et al. 2009. The STRFs of Ov neurons could be classified into three major types, namely NB and BB neurons (also observed in MLd and field L) and FS neurons (relatively abundant in Ov, although a few present in field L as well). Within each group, some tuning properties vary continuously, and these differences could have further significance in terms of coding sound features. For instance, in the BB group, the slower neurons are good at encoding the amplitude envelope, a sound feature important for timbre, while the faster BB neurons are good onset encoders, a sound feature important for following the rhythm of the song. In the NB group, both NB-S and the 2B neurons are selective for sounds with spectral modulations (i.e., succession of energy and lack of energy along the frequency axis), and thus could play a role in the recognition of sounds with a particular harmonic structure (see Woolley et al. 2009). Finally, because frequency sweeps are abundant in conspecific song, one might expect to find neurons in the avian auditory system to be tuned to sweeps, as has been demonstrated in the mammalian MGB (Lui and Mendelson 2003; Miller et al. 2002) and IC (Fuzessery et al. 2006; Hurley and Pollack 1999; Yue et al. 2007). In this study, we found that the avian thalamus exhibits a sensitivity to frequency sweeps not observed in MLd or to the same degree in field L: FS cells are first observed in the auditory thalamus and, perhaps somewhat surprisingly, appear less frequently and with lower modulation tuning at higher levels of the auditory system (i.e., a subset of HY cells in Woolley et al. 2009). Thus taken together, these various receptive fields found in the auditory thalamus are able to encode salient features of song.

It should be noted that the STRF does not capture all the tuning properties of auditory neurons (e.g., Gill et al. 2006; Linden et al. 2003; Woolley et al. 2009), such as sensitivity to context (e.g., Ahrens et al. 2008; Blake and Merzenich 2002; Nagel and Doupe 2008; Theunissen et al. 2000; Woolley et al. 2006) and to behavioral training (e.g., Fritz et al. 2003). In our analysis, our Ov STRFs captured ~42% of the response variance ($R^2 = 0.65^2$) of the neural response to conspecific song. Although this is far from a complete description, we believe that the linear STRF provides an accurate representation of the basic underlying computations performed by these cells in the sense that the more successful nonlinear models currently being investigated incorporate the linear STRF as a central piece of a more complex cascade of linear and nonlinear elements (Ahrens et al. 2008; Pillow et al. 2005) or perform a STRF on a nonlinear transformation of the stimulus (Gill et al. 2008). Moreover, the degree of the goodness of fit of the STRF model is similar in MLd and field L, which allowed us to use it for comparative purposes across brain regions. That said, it is clear that other response properties of Ov, such as sensitivity to context and to memory, have not been investigated here. These could reflect neural plasticity and might substantially differ in the auditory thalamus and auditory cortex.

Role of the auditory thalamus in transmitting auditory information from the midbrain to the pallium

The classic view of the primary sensory thalamic nuclei has been one of a simple relay center, where basic sensory information from the midbrain is relayed with fidelity to the cortex (Bilgake-Kunz et al. 1987; Sherman and Guillery 2002). Recent studies however have built a more complex and nuanced view of the sensory thalamus (Hu 2003; Sherman 2005, 2007). In the visual system, it is still thought that the primary thalamic nucleus of the lateral geniculate nucleus (LGN) relays similar receptive fields from retinal input to the visual cortex (in the

more classical sense), but this driver/relay information is modulated by cortical feedback to possibly improve the saliency of specific sensory stimuli or by modulatory nonretinal synapses related to behavioral state, including attention (Alitto and Usrey 2003; Kimura et al. 2005; Sherman 2001; Sherman and Guillery 2002; Zhang and Yan 2008). However, it is at the thalamo-cortical synapse that center-surround LGN receptive fields combine to form novel and distinct simple cell responses in V1 (see Hubel and Weisel 1998).

The emerging picture of the computations occurring at the midbrain, thalamic and cortical/pallial levels in the auditory system is a very different one compared with the visual system. First, studies investigating the transformation of the receptive fields at the *midbrain-thalamic synapse* in the mammalian auditory system have been sparse (see Eggermont 2001 review), and what is known comes mainly from sound localization studies. It is thought that the IC and the auditory thalamus are already quite sophisticated in processing sounds as revealed by the emergence of various space maps at these levels in both mammals and barn owls (Knudsen and Konishi 1978; Schnupp and King 1997; Yan and Suga 1996), which are further processed in higher processing centers. Two distinct pathways are implicated in processing auditory spatial signals in the avian brain: the forebrain (or pallial) pathway (IC – Ov – field L) and the midbrain pathway IC – ICx). The spatially tuned neurons in Ov: respond to localization cues without any apparent topographic mapping of these cues within the nucleus, in contrast to the midbrain pathway (Proctor and Konishi 1997); respond to lower frequencies and are more broadly tuned to interaural time differences (ITD) than in ICx (Perez and Pena 2006; Vonderschen and Wagner 2009); respond to a broader frequency range and their ITD and interaural level differences (ILD) tuning varies more across frequency (Perez et al. 2009) in contrast to the frequency independent ITD tuning of space-specific neurons in ICx neurons (Perez and Pena 2006). These differences found in the midbrain and pallial pathways demonstrate that all of the auditory midbrain (i.e., ICx) and thalamus do not appear to repeat exactly the same processing.

And unlike in the visual system with significant transformations occurring at the LGN-V1 synapse, studies in the mammalian auditory system have shown either incremental or moderate transformations across auditory *thalamocortical synapses*: there is a temporal slowing of responses in auditory cortex as compared with the auditory thalamus (Creutzfeldt et al. 1980); some properties and receptive fields are propagated with great fidelity while others are significantly transformed or generated intracortically (Miller et al. 2001); and finally, MGB neurons are able to synchronize to periodic clicks at repetition rates significantly higher than auditory cortex neurons, whereas nonsynchronized MGB responses typically occur at higher repetition rates than those observed in the auditory cortex (Bartlett and Wang 2007).

Few neurophysiological studies have probed Nucleus Ov in songbirds and studied its role in processing natural sounds [however, some early work can be found in Bigalke-Kunz et al. (1987); Durand et al. (1992); Ströhmann et al. (1994); and Lei and Mooney (2010) manipulated singing-related activity of feedback-sensitive Ov neurons and found that it triggered vocal plasticity]. In this study, we investigated whether the classic view of the first order relay station (where the receptive fields in the thalamus are similar to the ones in the midbrain as in the mammalian visual system) or whether the incremental feedforward model for the sensory thalamus (where distinct information is processed along the tecto-thalamo-telecephalic pathway), would hold true for the avian auditory thalamus processing vocalizations (as in the case of spatial receptive fields in barn owls). Recent work examining and comparing the range of response types at the level of the midbrain and the primary auditory pallium (Woolley et al. 2009) found that neurons in field L make a more heterogeneous group than those in MLd. On the other hand, some of the major functional groups found in MLd were also found in field L, indicating either a preservation or regeneration of similar tuning. Thus as auditory information progresses through the auditory processing stream, both the preservation of tuning and the generation of novel, more complex tuning occurs.

Here the STRFs derived from the auditory thalamus of zebra finches similarly reveal not only a preservation of the types of receptive fields inherited from the midbrain but also a generation of novel STRF types: the NB-S and FS neurons. We can therefore easily rule out the model that suggests that the first order sensory thalamus acts as a simple classic relay station (model A of Fig. 9). NB-S neurons in Ov can be interpreted as having intermediate tuning between the NB-T neurons found in MLd and the NB-S neurons in field L. Ov's NB-S neurons then allow for the possibility of the transformation schematized in model B in Fig. 9, where a linear scaling or sum of inputs

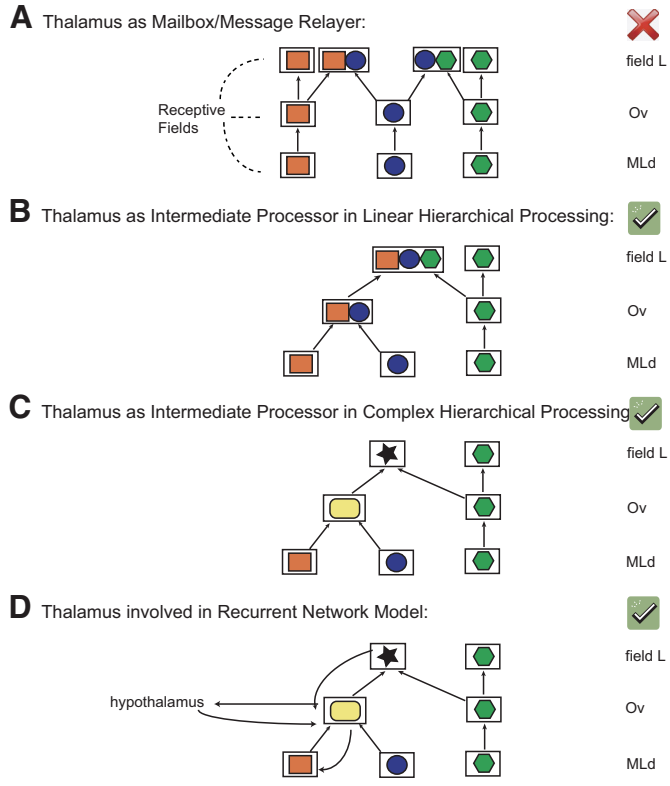

FIG. 9. Schematics of receptive fields (each denoted by a different symbol and color) at the midbrain, thalamic, and pallial levels, and 4 putative models of how these receptive fields are generated. *A*: classical view of the thalamus as a relay station in a hierarchical model. *B*: thalamus as an intermediate processing station in a linear processing chain. *C*: thalamus as an intermediate processing station in a complex hierarchical model. *D*: thalamus involved in a recurrent (feedforward and feedback) network model. Our data rejects Model A.

generates more complex receptive fields at each subsequent processing level. On the other hand, it is more difficult to demonstrate how FS neurons could be simple intermediates between neurons found in MLd and field L. FS neural tuning could emerge from a linear combination of NB neurons from MLd but with particular delays introduced to generate a sweep-sensitive cell. Or FS neurons (along with other complexly-tuned neurons) might result from circuitry that performs not only delayed linear operations but also nonlinear computations to generate more complex feature detectors, as schematized in model C in Fig. 9. In future studies, we plan on modeling Ov neurons using MLd STRFs as a basis set and similarly model field L neurons using Ov STRFs as a basis set. This analysis would allow us to determine more precisely the importance of delay lines and nonlinearities in the feed forward circuitry.

Our latency data (Fig. 8*A*) show the presence of some neurons with relatively long latencies, leaving open the possibility that hypothalamic input or cortical feedback plays a role even in the initial response to sounds as observed in the STRF. A more complete circuitry diagram for explaining these responses might therefore have to include a recurrent network architecture as schematized in model D in Fig. 9. In support of this model, we also found that earlier Ov responses are more similar to MLd STRFs (due to fast feedforward connections) and latter Ov responses are more similar to field L STRFs, allowing for the possibility of modulatory feedback from field L (see Fig. 8*B*). Anatomical studies have confirmed indirect feedback from field L via the HVC shelf, the RA (the robust-nucleus of the arcopallium) cup, and the Ov shell (see Fig. 1) that could further shape response properties of Ov neurons (Mello et al. 1998; Vates et al. 1996; Zeng et al. 2004). Cheng and Peng (1997) have also reported connections from the Ov shell to the anterior and posterior medial hypothalamus with reciprocal connections to Ov proper. Ov also receives input from the dorsal and ventrolateral lemniscal nuclei (shown in pigeons by Wild 1987 but not demonstrated in songbirds yet) that is also thought to mediate limbic function, although these secondary inputs from the hypothalamus and lemniscal nuclei are not as robust as those from the primary midbrain input of MLd. Thus even though it is not possible to rule out that the new types of STRFs observed in Ov emerge from projections from somewhere else than MLd, we hypothesize as in the visual system, that these secondary inputs perform more modulatory or limbic functions (i.e., memory) and that the primary midbrain input of MLd is driving the receptive field formation. However, this needs to be further tested, and obtaining receptive fields in the awake bird (and reversible inactivation studies) will further elucidate the direct role of feedback and limbic connections in shaping Ov responses.

Unique role for the avian auditory thalamus in processing vocalizations?

Given the differences in STRFs between Ov and field L (i.e., NB-S neurons with shorter temporal bandwidths) and the complexity observed in Ov (i.e., more pronounced and prevalent FS cells), an intriguing possibility arises where Ov plays a unique role in processing vocalizations independent of downstream auditory areas such as field L. For example, it is possible that as Ov processes a bird's call amid other environmental sounds, FS-type cells in Ov activate and send this information to the hypothalamus which in turn inhibits Ov proper to possibly block noncall or nonsong auditory inputs to enhance species-specific call effects (as suggested by Cheng and Peng 1997). Indeed the importance of Ov in processing certain behaviorally-relevant auditory information has been shown in the parrot, where the presentation of contact calls leads to a rapid expression of mRNA for the different subunits of the *N*-methyl-D-aspartate receptor in Ov (Brauth et al. 2007). Finally, Ov appears to be the main source of auditory information to the ventral paleostriatum (VP), which in turn might provide modulatory cholinergic input to the song nuclei HVC and RA, thereby indirectly influencing the processing of the bird's own song as well (Li et al. 1999; but also see Akutagawa and Konishi 2005). These findings are interesting but do not yet provide a definite answer as to whether Ov plays a unique role in processing vocalizations. Further experiments involving chronic recordings in different behavioral contexts and potentially with reversible inactivation of field L are required to further elucidate the function of Ov in this complex auditory circuit.

In summary, our data and analyses support a different view of the primary sensory thalamus than even the one regarded as the more nuanced one by the field in recent years. The emerging picture is that of the primary sensory thalamus playing a more substantial role than previously thought in building a high-level and somewhat distinct auditory representation of complex sounds from simpler acoustic features or components inherited from lower auditory centers (Winer et al. 2005). The avian auditory thalamus might also play a uniquer role in the recognition and the gating of behaviorally relevant sounds.

ACKNOWLEDGMENTS

We thank Y. M. for outstanding animal husbandry and B. R. for invaluable histological assistance.

GRANTS

This work was supported by an National Institute of Mental Health Grant MH-59189 to F. E. Theunissen and a National Aeronautic and Space Administration Graduate Research fellowship to N. Amin.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

Ahrens MB, Linden JF, Sahani M. Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. *J Neurosci* 28: 1929–1942, 2008.

Akutagawa E, Konishi M. Connections of thalamic modulatory centers to the vocal control system of the zebra finch. *Proc Natl Acad Sci USA* 102: 14086–14091, 2005.

Alitto HJ, Usrey WM. Corticothalamic feedback and sensory processing. *Curr Opin Neurobiol* 13: 440–445, 2003.

Bartlett EL, Wang X. Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *J Neurophysiol* 97: 1005–1017, 2007.

Bigalke-Kunz B, Rübsamen R, Dörrscheidt GJ. Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. *J Comp Physiol [A]* 161: 255–265, 1987.

Blake DT, Merzenich MM. Changes of A1 receptive fields with sound density. *J Neurophysiol* 88: 3409–3420, 2002.

Brauth SE, Liang W, Tang Y, Galdzicka E, Hall WS. Rapid contact call-driven induction of NR2A and NR2B NMDA subunit mRNAs in the

auditory thalamus of the budgerigar (*Melopsittacus undulatus*). *Neurobiol Learn Mem* 88: 33–39, 2007.

Capsius B, Leppelsack HJ. Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hear Res* 96: 59–70, 1996.

Cheng MF, Peng JP. Reciprocal talk between the auditory thalamus and the hypothalamus: an antidromic study. *Neuroreport* 8: 653–658, 1997.

Creutzfeldt O, Hellweg FC, Schreiner C. Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 39: 87–104, 1980.

Durand SE, Tepper JM, Cheng MF. The shell region of the nucleus ovoidalis: a subdivision of the avian auditory thalamus. *J Comp Neurol* 323: 495–518, 1992.

Eggermont JJ. Sound and perception: reviewing the search for a neural code. *Hear Res* 57: 1–42, 2001.

Feng AS, Narins PM, Capranica RR. Three populations of auditory fibers in the bullfrog. *J Comp Physiol [A]* 100: 221–229, 1975.

Fitzpatrick DC, Kanwal JS, Butman JA, Suga N. Combination-sensitive neurons in the primary auditory cortex of the mustached bat. *J Neurosci* 13: 931–940, 1993.

Fritz J, Shamma S, Elhilali M, Klein D. Rapid task-related plasticity of spectro-temporal receptive fields in primary auditory cortex. *Nat Neurosci* 6: 1216–1223, 2003.

Fuzessery ZM, Richardson MD, Coburn MS. Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the inferior colliculus of the pallid bat. *J Neurophysiol* 96: 1320–1336, 2006.

Gentner TQ, Margoliash D. Neuronal populations and single cells representing learned auditory objects. *Nature* 424: 669–674, 2003.

Gill P, Woolley SM, Fremouw T, Theunissen FE. What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. *J Neurophysiol* 99: 2809–2820, 2008.

Gill P, Zhang J, Woolley SM, Fremouw T, Theunissen FE. Sound representation methods for spectro-temporal receptive field estimation. *J Comput Neurosci* 21: 5–20, 2006.

Grace JA, Amin N, Singh NC, Theunissen FE. Selectivity for conspecific song in the zebra finch auditory forebrain. *J Neurophysiol* 89: 472–487, 2003.

Hsu A, Borst A, Theunissen FE. Quantifying variability in neural responses and its application for the validation of model predictions. *Network* 15: 91–109, 2004a.

Hsu A, Woolley SM, Fremouw TE, Theunissen FE. Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *J Neurosci* 24: 9201–9211, 2004b.

Hu B. Functional organization of lemniscal and nonlemniscal auditory thalamus. *Exp Brain Res* 153: 543–549, 2003.

Hubel DH, Wiesel TN. Early exploration of the visual cortex. *Neuron* 20: 401–412, 1998.

Hurley LM, Pollak GD. Serotonin differentially modulates responses to tones and frequency-modulated sweeps in the inferior colliculus. *J Neurosci* 19: 8071–8082, 1999.

Janata P, Margoliash D. Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J Neurosci* 19: 5108–5118, 1999.

Kelley DB, Nottebohm F. Projections of a telencephalic auditory nucleus-Field L in the canary. *J Comp Neurol* 183: 455–470, 1979.

Kemp DT. Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. *Eur Arch Otorhinolaryngol* 224: 37–45, 1979.

Kimura A, Donishi T, Okamoto K, Tamai Y. Topography of projections from the primary and non-primary auditory cortical areas to the medial geniculate body and thalamic reticular nucleus in the rat. *Neuroscience* 135: 1325–1342, 2005.

Knudsen EI, Konishi M. A neural map of auditory space in the owl. *Science* 200: 795–797, 1978.

Kowalski N, Depireux DA, Shamma SA. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J Neurophysiol* 76: 3503–3523, 1996.

Langner G, Bonke D, Scheich H. Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. *Exp Brain Res* 43: 11–24, 1981.

Lei H, Mooney R. Manipulation of a central auditory representation shapes learned vocal output. *Neuron* 65: 122–134, 2010.

Leppelsack HJ, Vogt M. Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *J Comp Neurol* 107: 263–274, 1976.

Lewicki MS, Arthur BJ. Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16: 6987–6998, 1996.

Li R, Zuo MX, Sakaguchi H. Auditory-vocal cholinergic pathway in zebra finch brain. *Neuroreport* 10: 165–169, 1999.

Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* 90: 2660–2675, 2003.

Liu RC, Schreiner CE. Auditory cortical detection and discrimination correlates with communicative significance. *PLoS Biol* 5(7): e173, 2007.

Lui B, Mendelson JR. Frequency modulated sweep responses in the medial geniculate nucleus. *Exp Brain Res* 153: 550–553, 2003.

Margoliash D, Fortune ES. Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. *J Neurosci* 12: 4309–4326, 1992.

Miller LM, Escabí MA, Read HL, Schreiner CE. Functional convergence of response properties in the auditory thalamocortical system. *Neuron* 32: 151–160, 2001.

Miller LM, Escabi MA, Read HL, Schreiner CE. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J Neurophysiol* 87: 516–527, 2002.

Nagel KI, Doupe AJ. Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. *Neuron* 58: 938–955, 2008.

Newman J, Wollberg Z. Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. *Brain Res* 54: 287–304, 1978.

Ohl FW, Scheich H, Freeman WJ. Change in pattern of ongoing cortical activity with auditory category learning. *Nature* 412: 733–736, 2001.

Perez ML, Pena JL. Comparison of midbrain and thalamic space-specific neurons in barn owls. *J Neurophysiol* 95: 783–790, 2006.

Perez ML, Shanbhag SJ, Pena JL. Auditory spatial tuning at the crossroads of the midbrain and forebrain. *J Neurophysiol* 102: 1472–1482, 2009.

Pillow JW, Paninski L, Uzzell VJ, Simoncelli EP, Chichilnisky EJ. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J Neurosci* 25: 11003–11013, 2005.

Plummer TK, Striedter GF. Auditory responses in the vocal motor system of budgerigars. *J Neurobiol* 42: 79–94, 2000.

Pollak GD, Klug A, Bauer EE. Processing and representation of species-specific communication calls in the auditory system of bats. *Int Rev Neurobiol* 56: 83–121, 2003.

Proctor L, Konishi M. Representation of sound localization cues in the audiotry thalamus of the barn owl. *Proc Natl Acad Sci USA* 94: 10421–10425, 1997.

Qiu A, Schreiner CE, Escabí MA. Gabor analysis of auditory midbrain receptive fields: spectro-temporal and binaural composition. *J Neurophysiol* 90: 456–476, 2003.

Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268: 111–114, 1995.

Scheich H, Langner G, Koch R. Coding of narrow-band and wide-band vocalizations in the auditory midbrain nucleus (MLD) of the guinea fowl (*Numida meleagris*). *J Comp Physiol* 117: 245–265, 1977.

Schnupp JW, King AJ. Coding for auditory space in the nucleus of the brachium of the inferior colliculus in the ferret. *J Neurophysiol* 78: 2717–2731, 1997.

Sen K, Theunissen FE, Doupe AJ. Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol* 86: 1445–1458, 2001.

Sherman SM, Guillery RW. The role of the thalamus in the flow of information to the cortex. *Philos Trans R Soc Lond B Biol Sci* 357: 1695–1708, 2002.

Sherman SM. Thalamic relay functions. *Prog Brain Res* 134: 51–69, 2001.

Sherman SM. Thalamic relays and cortical functioning. *Prog Brain Res* 149: 107–126, 2005.

Sherman SM. The thalamus is more than just a relay. *Curr Opin Neurobiol* 17: 417–422, 2007.

Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *JASA* 114: 3394–3411, 2003.

Ströhmann B, Schwarz DW, Puil E. Mode of firing and rectifying properties of nucleus ovoidalis neurons in the avian auditory thalamus. *J Neurophysiol* 71: 1351–1360, 1994.

Theunissen FE, David SV, Singh NC, Hsu A, Vinje W, Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network Comp Neural Syst* 12: 1–28, 2001.

Theunissen FE, Sen K, Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20: 2315–2331, 2000.

Theunissen FE, Woolley SM, Hsu A, Fremouw T. Methods for the analysis of auditory processing in the brain. *Ann NY Acad Sci* 1016: 187–207, 2004.

Ulanovsky N, Las L, Nelken I. Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6: 391–398, 2003.

Vates GE, Broome BM, Mello CV, Nottebohm F. Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taenopygia guttata*). *J Comp Neurol* 366: 613–642, 1996.

Vonderschen K, Wagner H. Tuning to interaural time difference and frequency differs between the auditory arcopallium and the external nucleus of the inferior colliculus. *J Neurophysiol* 101: 2348–2361, 2009.

Wang X, Merzenich M, Beitel R, Schreiner C. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74: 2685–2706, 1995.

Wild JM. Nuclei of the lateral lemniscus project directly to the thalamic auditory nuclei in the pigeon. *Brain Res* 408: 203–207, 1987.

Winer JA, Miller LM, Lee CC, Schreiner CE. Auditory thalamocortical transformation: structure and function. *Trends Neurosci* 28: 255–263, 2005.

Woolley SM, Fremouw TE, Hsu A, Theunissen FE. Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat Neurosci* 8: 1371–1379, 2005.

Woolley SM, Gill PR, Fremouw T, Theunissen FE. Functional groups in the avian auditory system. *J Neurosci* 29: 2780–2793, 2009.

Woolley SM, Gill PR, Theunissen FE. Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *J Neurosci* 26: 2499–2512, 2006.

Xie R, Meitzen J, Pollak GD. Differing roles of inhibition in hierarchical processing of species-specific calls in auditory brainstem nuclei. *J Neurophysiol* 94: 4019–4037, 2005.

Yan J, Suga N. The midbrain creates and the thalamus sharpens echo-delay tuning for the cortical representation of target-distance information in the mustached bat. *Hear Res* 93: 102–110, 1996.

Yue Q, Casseday JH, Covey E. Response properties and location of neurons selective for sinusoidal frequency modulations in the inferior colliculus of the big brown bat. *J Neurophysiol* 98: 1364–1373, 2007.

Zeng S, Zhang X, Peng W, Zuo M. Immunohistochemistry and neural connectivity of the Ov shell in the songbird and their evolutionary implications. *J Comp Neurol* 470: 192–209, 2004.

Zhang Y, Yan J. Corticothalamic feedback for sound-specific plasticity of auditory thalamic neurons elicited by tones paired with basal forebrain stimulation. *Cereb Cortex* 18: 1521–1528, 2008.