

# Binaural Gain Modulation of Spectrotemporal Tuning in the Interaural Level Difference-Coding Pathway

Louisa J. Steinberg,<sup>1</sup> Brian J. Fischer,<sup>2</sup> and Jose L. Peña<sup>1</sup>

<sup>1</sup>Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461 and <sup>2</sup>Department of Mathematics, Seattle University, Seattle, Washington 98122

In the brainstem, the auditory system diverges into two pathways that process different sound localization cues, interaural time differences (ITDs) and level differences (ILDs). We investigated the site where ILD is detected in the auditory system of barn owls, the posterior part of the lateral lemniscus (LLDp). This structure is equivalent to the lateral superior olive in mammals. The LLDp is unique in that it is the first place of binaural convergence in the brainstem where monaural excitatory and inhibitory inputs converge. Using binaurally uncorrelated noise and a generalized linear model, we were able to estimate the spectrotemporal tuning of excitatory and inhibitory inputs to these cells. We show that the response of LLDp neurons is highly locked to the stimulus envelope. Our data demonstrate that spectrotemporally tuned, temporally delayed inhibition enhances the reliability of envelope locking by modulating the gain of LLDp neurons' responses. The dependence of gain modulation on ILD shown here constitutes a means for space-dependent coding of stimulus identity by the initial stages of the auditory pathway.

## Introduction

The brainstem auditory system, an obligatory pathway for all aural information, must encode the location of a sound source, as well as its identity. Encoding information about spectrotemporal features is essential for discriminating between sounds, such as in speech and species-specific vocalizations (Shannon et al., 1995; Smith et al., 2002; Woolley et al., 2005; Suta et al., 2007; Nelson and Takahashi, 2010; Schneider and Woolley, 2010; Fogerty, 2011; Wang et al., 2011). In the auditory pathway, the timing of the spikes carries spectrotemporal information (Hermes et al., 1981; Theunissen et al., 2000; Escabí and Read, 2003; Linden et al., 2003; Joris et al., 2004; Christianson and Peña, 2007; Steinberg and Peña, 2011). This putative temporal code must be preserved with fidelity over several processing stages.

Barn owls localize sound sources using interaural time differences (ITDs) and interaural level differences (ILDs) (Konishi, 2003). In mammals and birds, these cues are processed in two parallel brainstem pathways (Schnupp and Carr 2009). In the barn owl, auditory nerve fibers bifurcate, one branch projecting to the cochlear nucleus angularis (NA) and the other to the cochlear nucleus magnocellularis (NM), giving rise to the ILD- and ITD-processing pathways, respectively (Sullivan and Konishi, 1984; Takahashi et al., 1984; Carr and Boudreau, 1991). Compared to NM neurons, NA neurons show en-

hanced ability to encode the stimulus envelope through increased sensitivity to power transients within their preferred frequency range (Steinberg and Peña, 2011; Kreeger et al. 2012). Thus, at the starting point of the ILD- and ITD-processing pathways, the former encodes the stimulus envelope with higher fidelity than the latter.

The posterior part of the lateral lemniscus (LLDp) is the first binaural nucleus in the ILD-processing pathway. Similar to cells in the mammalian lateral superior olive, cells here detect ILDs through the interplay of excitation and inhibition (Galambos et al., 1959; Boudreau and Tsuchitani, 1968). LLDp cells are excited by contralateral stimuli via NA and inhibited by ipsilateral stimuli via reciprocal connections from the LLDp of the opposite hemisphere (Manley et al., 1988; Takahashi and Konishi, 1988; Takahashi and Keller, 1992; Mogdans and Knudsen, 1994; Takahashi et al., 1995). This interplay of excitation and inhibition gives rise to sigmoid ILD tuning curves, where cells respond maximally to sounds that are louder in the contralateral ear.

Here we show how excitation and inhibition interact to enhance the fidelity of envelope locking in LLDp. Using binaurally uncorrelated noise, we measured the spectrotemporal receptive fields (STRFs) of the excitatory and inhibitory monaural inputs that converge onto LLDp neurons. We show that the balance of excitation and inhibition, which changes with sound location, modulates the gain of these neurons. This leads to greater response reliability of LLDp neurons to the stimulus envelope, meaning the reproducibility of precise spike timing across repeated presentations of a unique stimulus, compared to the last processing stage of the ITD pathway. This finding lends support to the hypothesis that the ILD pathway transmits spectrotemporal information to downstream structures with greater fidelity than the ITD pathway. In addition, our work suggests that sound source location may play a role in coding stimulus identity.

Received Oct. 22, 2012; revised May 2, 2013; accepted May 22, 2013.

Author contributions: L.J.S. and J.L.P. designed research; L.J.S. performed research; L.J.S. contributed unpublished reagents/analytic tools; L.J.S. and B.J.F. analyzed data; L.J.S., B.J.F., and J.L.P. wrote the paper.

This work was supported by National Institute of Health Grant DC007690. We are grateful to Bertrand Fontaine for comments on this manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Louisa J. Steinberg, Rose F. Kennedy Center, Room 529, 1410 Pelham Parkway South, Bronx, NY 10461. E-mail: louisa.steinberg@med.einstein.yu.edu.

DOI:10.1523/JNEUROSCI.4941-12.2013

Copyright © 2013 the authors 0270-6474/13/3311089-11\$15.00/0

## Materials and Methods

**Surgery.** Data were collected from five adult barn owls (*Tyto alba*) bred in captivity. Since it is not possible to verify an owl's sex without invasive methods, sexes were undetermined. As described previously by Steinberg and Peña (2011), the birds were anesthetized by intramuscular injection of ketamine hydrochloride (20 mg/kg; Ketaset) and xylazine (4 mg/kg; Anased) over the course of the experiment. The depth of anesthesia was monitored by toe pinch. They also received an intramuscular injection of prophylactic antibiotics (oxytetracycline; 20 mg/kg; Phoenix Pharmaceuticals) and a subcutaneous injection of lactated Ringers solution (10 ml) at the beginning for each experiment. Body temperature was maintained throughout the experiment with a heating pad (American Medical Systems). A metal head plate was implanted at the beginning of the first recording session by removing the top layer of the skull and affixing it with dental cement while the head was held in stereotaxic position with ear bars and a beak holder. A small steel post was also implanted to demarcate a reference point for stereotaxic coordinates. Subsequently, all recording sessions were performed while the head was held in place by the head plate. A well was created on the skull around the stereotaxic coordinates for the LLDp, using dental cement, and the skin was sutured around it. A craniotomy was performed at the coordinates for the recording site, and a small incision was made in the dura mater for electrode insertion. At the end of a recording session, the craniotomy was sealed with Rolyan silicone elastomer (Sammons Preston). After the experiment, analgesics (ketoprofen, 1 mg/kg; Ketofen; Fort Dodge) were administered. Owls were returned to individual cages and monitored for recovery. These procedures comply with guidelines set forth by the National Institutes of Health and the Albert Einstein College of Medicine's Institute of Animal Studies.

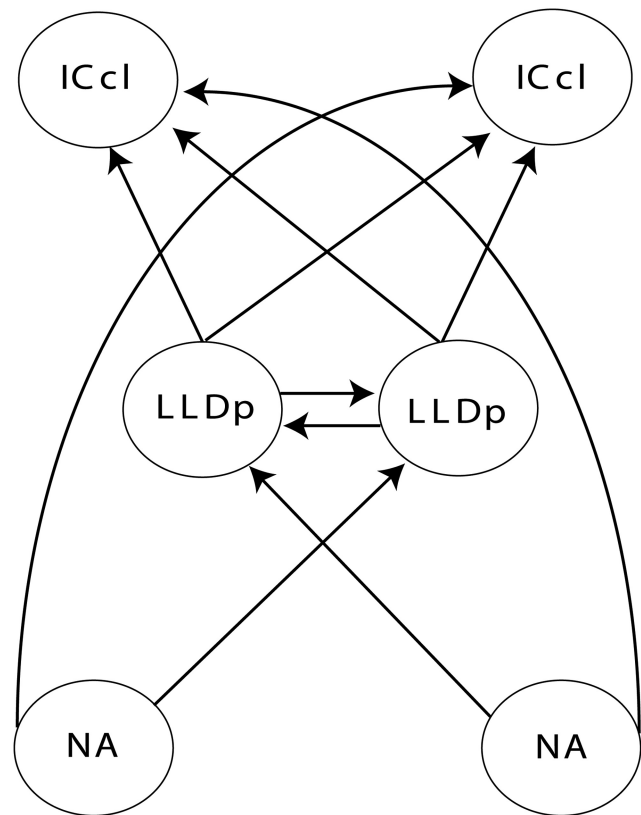
**Acoustic stimulation.** Dichotic stimulation was delivered in a double-walled sound-attenuating chamber (Industrial Acoustics). Custom software was used to generate stimuli and collect data. Earphones were constructed from a small speaker (model 1914; Knowles) and a microphone (model 1319; Knowles) in a custom-made aluminum case that fit the owl's ear canal. The microphones were calibrated using a Brüel and Kjaer microphone, allowing us to translate voltage output into decibels of SPL. The calibrated microphones were used to calibrate the earphones at the beginning of each experiment while inside the owl's ear canal. The calibration data contained the amplitudes and phase angles measured in frequency steps of 100 Hz. The stimulus generation software then used these calibration data to automatically correct irregularities in the amplitude and phase response of each earphone from 0.5 to 12 kHz (Arthur, 2004).

Acoustic stimuli consisted of broadband noise bursts with a linear rise and fall time of 5 ms.

**Electrophysiology.** The LLDp was targeted by stereotaxic coordinates. Units were identified based on response properties unique to LLDp, which receives excitatory input from the contralateral NA and reciprocal inhibition from the contralateral LLDp (Fig. 1). Extracellular spikes were recorded from LLDp neurons using 1 M $\Omega$  tungsten electrodes (A-M Systems) and amplified by a DP-301 differential amplifier (Warner Instruments).

A spike discriminator (SD1; Tucker-Davis Technologies) converted neural impulses into TTL pulses for an event timer (ET1; Tucker-Davis Technologies), which recorded the timing of the pulses.

**Data collection.** For each neuron we obtained rate-ITD, rate-ILD, and frequency tuning curves, as well as the monaural rate-intensity (RI) curve. ITD and ILD tuning curves were measured using 10 repetitions of 100-ms-long white-noise stimuli at randomly interleaved ITDs and ILDs. Frequency tuning curves were measured using 10 repetitions of 100 ms randomly interleaved pure tones between 500 Hz and 12 kHz presented at each neuron's best ILD. To estimate neurons' monaural RI curve, 10 repetitions of 100-ms-long white noise at sound levels from below threshold to saturation were presented in random order to the contralateral (excitatory) ear of each neuron (Table 1). To determine whether or not a neuron received inhibitory input from the ipsilateral ear, as opposed to just being excited by the contralateral ear, we performed the ILDF protocol, in which sound level is held fixed in the con-



**Figure 1.** Schematic depiction of the LLD-processing pathway with bilateral projections. The cochlear NA projects to the contralateral LLDp, as well as the contralateral ICcl. The LLDp nuclei of the two hemispheres reciprocally project to each other. LLDp also projects bilaterally to ICcl.

tralateral (excitatory) ear while varying the sound level in the ipsilateral (inhibitory) ear. Because LLDp neurons receive inhibitory input from the ipsilateral side, their response is depressed when the sound level in the ipsilateral ear reaches sufficient intensity (Manley et al., 1988). This protocol was performed with 100-ms-long noise bursts.

Data used to determine the neurons' STRFs and their response reliability were collected as described previously (Christianson and Peña, 2007; Steinberg and Peña, 2011). Briefly, to measure binaural STRFs, we presented a string of *de novo*-synthesized broadband noise (unfrozen noise protocol). To measure neurons' monaural STRFs, binaurally uncorrelated noise stimuli were presented (Fischer et al., 2008). To measure response reliability, we used a "frozen noise" protocol, in which a single broadband noise stimulus (1 to 12 kHz) was repeated at different combinations of ILDs and sound levels (Table 1). Frozen noise stimuli were generated *de novo* each time this stimulus protocol was run, such that frozen noise tokens are unique for different instances of this protocol even for the same neuron. In all cases, noise segments were 500 ms long with a rise and fall time of 5 ms and at a stimulation rate of 1 stimulus per second or less.

**Generalized linear model.** We used a generalized linear model (GLM) to describe spiking responses in LLDp (Paninski, 2004; Calabrese et al., 2011; Plourde et al., 2011; Fig. 2).

The response of the neuron is determined by the conditional intensity function  $\lambda(t|\mathbf{R}(t))$ , where  $\text{Prob}(\text{spike in } [t, t + \Delta]) \approx \lambda(t|\mathbf{R}(t))\Delta$  for small time bins  $\Delta$ .

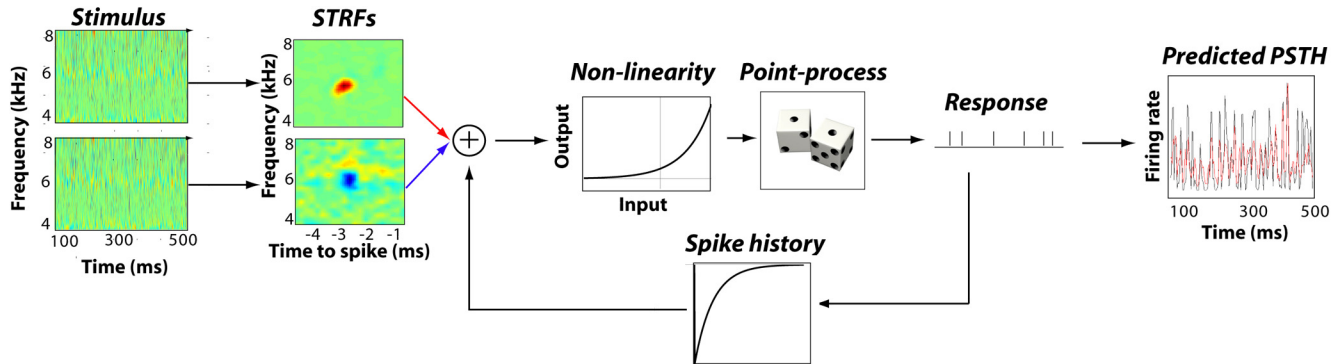
The response of the neuron  $r(t)$  equals one if there is a spike and zero otherwise. The spike history,  $\mathbf{R}(t)$ , consists of the response at the previous  $K$  time steps:  $\mathbf{R}(t) = [r(t - \Delta), r(t - 2\Delta), \dots, r(t - K\Delta)]$ .

The conditional intensity function is modeled such that spiking depends on (1) STRFs  $\mathbf{k}_L$  and  $\mathbf{k}_R$  applied to the left and right ear inputs  $\mathbf{s}_L(t)$  and  $\mathbf{s}_R(t)$ , and (2) a spike history term  $\mathbf{k}_{\text{spike}}$  that reflects refractory processes:  $\lambda[t|\mathbf{R}(t)] = \exp[\mathbf{k}_L \cdot \mathbf{s}_L(t) + \mathbf{k}_R \cdot \mathbf{s}_R(t) + \mathbf{k}_{\text{spike}} \cdot \mathbf{R}(t) + k_0] = \exp[\mathbf{k}_L \cdot \mathbf{s}_L(t)] \exp[\mathbf{k}_R \cdot \mathbf{s}_R(t)] \exp[\mathbf{k}_{\text{spike}} \cdot \mathbf{R}(t) + k_0]$ .

**Table 1. Overview of stimulation protocols**

Stimulus protocol	Binaural vs monaural	Binaural correlation	Noise type	Sound level/ABI	ILD	Stimulus duration
ILD tuning curve	Binaural	100%	Unfrozen	Constant (40 dB)	−40 to 40 dB	100 ms
ILDf	Binaural	100%	Unfrozen	Varying	Varying	100 ms
Monaural intensity	Monaural	N/A	Unfrozen	10 dB to 80 dB	N/A	100 ms
STRF estimation	Binaural	0%	Unfrozen	Constant (40 dB)	2 ILDs over dynamic range	500 ms
Response reliability, ILD protocol	Binaural	100%	Frozen	Constant (40 dB)	2–3 ILDs over dynamic range	500 ms
Response reliability, ILDf protocol	Binaural	100%	Frozen	Varying	Varying	500 ms
Response reliability, monaural excitation protocol	Monaural	N/A	Frozen	2–3 dBs in dynamic range	N/A	500 ms

N/A, Not applicable.



**Figure 2.** Schematic description of the GLM model. Each ear's input is given by a spectrotemporal representation of the sound signal. Each ear's STRF is applied to the stimulus to give a weighted sum over frequency and time. The spike history kernel is a temporal filter applied to the spiking history to take into account refractory processes. The STRF-filtered stimuli are combined with the spike history and passed through a nonlinearity to produce an instantaneous firing rate. Spikes generated probabilistically with the instantaneous firing rate give rise to a predicted PSTH (red), which is compared to the neuron's response observed *in vivo* (black).

Each ear's input is given by a spectrotemporal representation of the sound signal. The spectrotemporal representation of the sound is found by filtering the sound signal with a gammatone filter bank, computing the envelope of the output of each gammatone filter using the Hilbert transform, and taking the logarithm of the envelopes.

Each STRF is applied to one ear's input to give a weighted sum over frequency and time of the spectrotemporal input signal:

$$\mathbf{k}_R \cdot \mathbf{s}_R(t) = \sum_{n=1}^N \sum_{m=0}^M \mathbf{k}_R(n,m) \mathbf{s}_R(f_n, t - m\Delta).$$

The spike history kernel is a temporal filter applied to the spiking history:

$$\mathbf{k}_{\text{spike}} \cdot \mathbf{R}(t) = \sum_{k=1}^K \mathbf{k}_{\text{spike}}(k) r(t - k\Delta).$$

We used regularization to avoid overfitting. Regularization was incorporated into the model fitting procedure by adding a quadratic penalty on the model parameters to the log-likelihood of the data. Specifically, model parameters were found by maximizing the cost equal to  $L(\mathbf{k}) - Q(\mathbf{k})$ , where

$$L(\mathbf{k}) = \sum_{t_{\text{spk}}} \log[\lambda(t_{\text{spk}} | \mathbf{R}(t))] - \int \lambda(t | \mathbf{R}(t)) dt$$

is the log-likelihood of the data, and the quadratic penalty is given by the following:

$$Q(k) = \eta_1 \sum_{n=1}^N \sum_{m=0}^M [\mathbf{k}_L(n,m)]^2 + \eta_2 \sum_{n=1}^N \sum_{m=0}^M [\mathbf{k}_R(n,m)]^2 + \eta_3 \sum_{k=1}^K [\mathbf{k}_{\text{spike}}(k)]^2 + \eta_4 k_0^2.$$

The Matlab function `glmfitqp` (Mineault; Matlab Central) was used to find the parameters of the model that maximize the penalized log-likelihood of the observed spike response, binned at 0.5 ms.

We selected the scale parameters for the quadratic penalty  $\eta_i$  and the time and frequency ranges  $N$  and  $M$  for each neuron using cross-validation. We used responses to frozen, binaurally correlated noise bursts for cross-validation. In six neurons where the response to frozen noise was unavailable to use for cross-validation, we used a fivefold cross-validation. The cross-validation method tests the model by successively holding out a portion of the training data, fitting the model with the remaining data, and testing the model on the withheld data. This was done for five segments of the data set. A model peristimulus time histogram (PSTH) was generated over 10,000 trials, binned at 0.5 ms, and smoothed with a 5 ms Hanning window (Calabrese et al. 2011). The correlation coefficient between the model PSTH and the measured PSTH was used to assess the accuracy of the model.

**STRF properties.** STRFs were described by the best frequency, frequency bandwidth, latency, and temporal bandwidth. The best frequency and latency were computed as the frequency and time, respectively, at the peak of the STRF. The frequency bandwidth was computed as the width at half-height of the frequency slice through the STRF at the latency. Similarly, the temporal bandwidth was computed as the width at half-height of the temporal slice through the STRF at the best frequency.

**Characterizing gain modulation.** We characterized the gain modulation in the GLM output using responses of the model at two ILDs, one more inhibitory than the other, using the same frozen noise stimulus. The firing rate of the GLM is determined by the conditional intensity function  $f[x(t)]$ , where the input to the model is  $x(t) = \mathbf{k}_L \cdot \mathbf{s}_L(t) + \mathbf{k}_R \cdot \mathbf{s}_R(t) + \mathbf{k}_{\text{spike}} \cdot \mathbf{R}(t)$ , and the STRFs and spike history kernel depend on ILD. We presented a frozen noise stimulus to the models at each ILD and measured the input signal  $x(t)$ , averaged over trials, and the PSTH. We used linear regression to predict the input signal at the more inhibitory ILD  $x_1(t)$  from the input signal at the less inhibitory ILD  $x_2(t)$ , as  $ax_2(t) + b$ . We then predicted the PSTH of the GLM at the inhibitory ILD using the input signal at the more excitatory ILD and the parameters of

the linear regression as  $f(ax_2(t) + b)$ . Because the input–output nonlinearity is the exponential function, this prediction can be rewritten as  $f[ax_2(t) + b] = \exp[ax_2(t) + b] = \exp(b)\exp[ax_2(t)] = g_1 f[g_2 x_2(t)]$ , where the gain parameters  $g_1$  and  $g_2$  depend on ILD, but not on the stimulus. Therefore, the GLM response at the small ILD is predicted by applying a gain to the input signal at the large ILD and to the input–output function of the GLM.

**Response reliability.** The first 100 ms of each trial were excluded for the shuffled autocorrelogram (SAC), considering only those spikes that occurred once the firing rate had reached a steady state by visual inspection of the PSTH. In most neurons, this steady state was reached well before 100 ms after the stimulus onset in both frozen and unfrozen noise protocols.

Response reliability to the stimulus envelope was assessed by quantifying the reproducibility of the neural response to repeated presentations of the same stimulus using the SAC (Joris, 2003; Joris et al., 2006; Christianson and Peña, 2007; Steinberg and Peña, 2011). A spike train recorded during one presentation of the frozen noise stimulus is compared to all other spike trains recorded from other presentations of the same stimulus. For each possible combination, the forward time intervals are computed between pairs of spike trains using a 50  $\mu$ s bin width. A normalizing factor was used,  $N(N-1)r^2\Delta\tau D$ , where  $r$  is the mean firing rate,  $\Delta\tau$  is the bin width of the correlogram, and  $D$  is stimulus duration. This produces a unity baseline, where a spike train with Poisson statistics will have a flat SAC of height 1. The main parameter considered here to quantify neurons' response reliability is the height of this normalized SAC's peak at zero delay measured in number of normalized coincidences. This measure has also been referred to as the coincidence index.

To study how inhibitory drive influences response reliability, we used changes in firing rate as a proxy for changes in inhibitory input. We used three stimulation protocols: (1) the ILD protocol in which the average binaural intensity (ABI) was held constant while sounds were presented at two to three ILDs, (2) the ILDf protocol where sound level in the ipsilateral (inhibitory) ear was varied while keeping the sound level fixed at the contralateral (excitatory) ear, and (3) the monaural protocol in which sound level was varied in the contralateral (excitatory) ear without any sound presented to the ipsilateral (inhibitory) ear (Table 1). The combinations of ILDs, ABIs, and monaural intensity were chosen based on the dynamic range of each neuron. For the ILD and ILDf protocols, these values were chosen such that for one combination the neuron received minimum inhibitory drive. An additional one to two conditions were presented in which the neuron received a stronger inhibitory drive, but still produced enough spikes for SAC calculation. For the monaural stimulus protocol, intensities were presented such that a maximum firing rate was elicited, as well as two intensities around the center of the dynamic range, and another that elicited a lower firing rate. Since there is variation from neuron to neuron in spontaneous firing rate and dynamic range between stimulus protocols, we normalized changes in firing rate and response reliability (SAC peak height) for each neuron within each stimulus protocol to be able to compare across neurons. This resulted in what we call here the firing rate ratio ( $FR_{rat}$ ) and SAC ratio ( $SAC_{rat}$ ), defined as  $FR_{rat} = FR_{ILD_i}/FR_{ILD_{ex}}$  and  $SAC_{rat} = SAC_{ILD_i}/SAC_{ILD_{ex}}$ , where  $FR_{ILD_{ex}}$  and  $SAC_{ILD_{ex}}$  are the firing rate and SAC peak height, respectively, of conditions in the stimulus protocol at which the cell receives minimum inhibitory drive, and  $FR_{ILD_i}$  and  $SAC_{ILD_i}$  are the firing rate and SAC peak height of stimulus protocol conditions at which inhibitory drive is stronger. A similar calculation was conducted for the ILDf protocol, as well as the monaural stimulation protocol. We also repeated these calculations using the maximum firing rate across all stimulation protocols and its corresponding SAC peak as the denominator for  $FR_{rat}$  and  $SAC_{rat}$ , as described above, as a control for biases introduced within conditions. Data were fit using a curve-fitting tool, the function *cf*tool, in Matlab. *Cf*tool performs several goodness-of-fit assessments, including sum of squares due to error,  $R_2$ , and root mean squared error. These measures were used to assess which was the best function form used to fit the data.

Linear regressions derived from these data were compared using a  $t$  test for the comparison of two slopes, as described by Zar (2009). This analysis algorithm is shown below, where the subscripts 1 and 2 denote

data points from different data sets (in this case, two stimulus protocols, i.e., monaural vs binaural stimulation), and  $x$  and  $y$  denote  $FR_{rat}$  and  $SAC_{rat}$  data, respectively:

$$t = \frac{b_1 - b_2}{S_{b_1 - b_2}}$$

such that

$$b = \frac{\sum xy}{\sum x^2}$$

and

$$S_{b_1 - b_2} = \sqrt{\frac{(S_{Y^*X})_p}{\left(\sum x^2\right)_1} + \frac{(S_{Y^*X})_p}{\left(\sum x^2\right)_2}}$$

where

$$(S_{Y^*X})_p = \frac{(\text{residual SS})_1 + (\text{residual SS})_2}{(\text{residual DF})_1 + (\text{residual DF})_2}$$

Here, residual sum of squares (residual SS) and residual degrees of freedom (residual DF) are defined as follows:

$$\text{residual SS} = \sum y^2 - \frac{\left(\sum xy\right)^2}{\sum x^2},$$

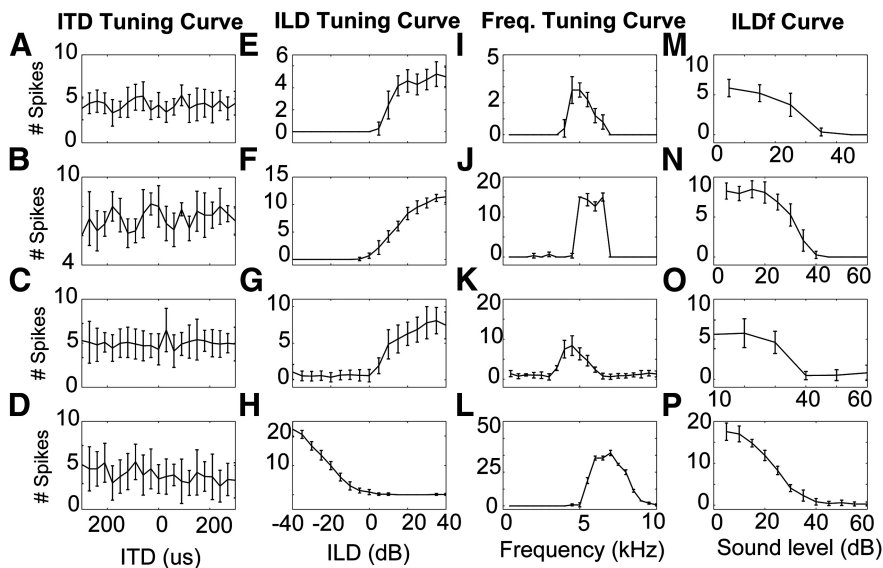
and residual DF =  $n - 2$ , where  $n$  is the number of data points for each stimulus protocol.

## Results

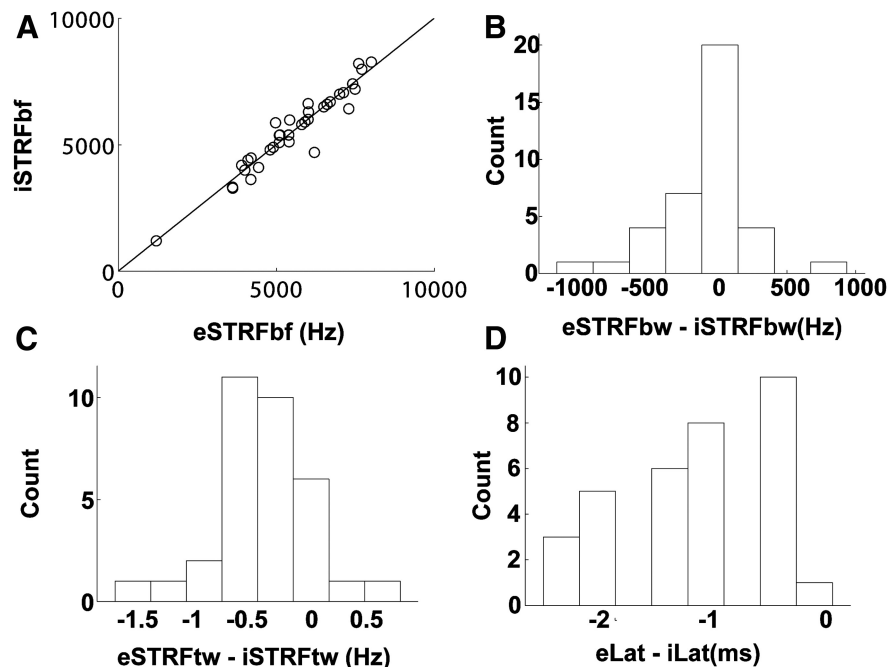
Our data set was obtained from 81 LLDp neurons recorded in five barn owls. LLDp neurons were targeted by stereotaxic coordinates and identified based on their response properties. LLDp neurons are unique within the auditory brainstem of the barn owl in that they are excited exclusively by sound in the contralateral ear, inhibited by the ipsilateral ear, narrowly tuned to frequency, and not tuned to ITD (Manley et al., 1988; Mogdans and Knudsen, 1994). To classify a neuron as belonging to the LLDp, all of these response properties were confirmed (Fig. 3). Therefore, LLDp cells can be distinguished from the cochlear nuclei, which respond to ipsilateral stimuli and from all other binaural nuclei in the region, tuned to ITD. Specifically, LLDp is located anatomically close to the anterior part of the dorsal lateral lemniscus, which exhibits robust ITD tuning; this prevents confusion of the two. Within the LLDp, there is an anatomical gradient of inhibitory input from dorsal to ventral, such that cells in the dorsal part of the nucleus receive significant inhibition, whereas cells in the most ventral part of the nucleus receive little inhibition (Manley et al., 1988; Takahashi et al., 1989; Takahashi and Keller, 1992). Thus, most LLDp cells are inhibited when sound is increased in the ipsilateral ear while holding the sound level constant in the contralateral (excitatory) ear.

### STRFs of LLDp cells

We estimated the STRFs of 40 LLDp neurons using a point process model of LLDp spiking responses. As demonstrated previously in the barn owl's nucleus laminaris (Fischer et al., 2008; Fischer et al., 2011), binaurally uncorrelated noise can be used to study the monaural inputs of auditory neurons. We used this strategy in LLDp neurons to estimate the STRFs of the excitatory, contralateral (eSTRF) and the inhibitory, ipsilateral (iSTRF) monaural inputs. Unlike the tuning to ITD in the nucleus lami-



**Figure 3.** Example of response properties of four LLDp neurons. Each row represents one neuron. **A–D**, Rate-ITD curves. (These neurons are not tuned to ITD.) **E–H**, Rate-ILD curves. **I–L**, Isointensity frequency-tuning curves. **M–P**, Rate-ILDf curves, where sound level was held constant in the contralateral (excitatory) ear while it was varied in the ipsilateral (inhibitory) ear as indicated on the *x*-axis. Neurons in the first three rows were recorded from the left hemisphere, while the neuron recorded in the last row was recorded from the right hemisphere. Therefore, the ILD curve in the last row is reversed relative to the other neurons. Error bars indicate SEM.



**Figure 4.** Comparison of monaural excitatory and inhibitory STRF properties. **A**,  $eSTRF_{bf}$  and  $iSTRF_{bf}$  were strongly correlated ( $iSTRF_{bf} = eSTRF_{bf} + 7$  Hz,  $r = 0.96$ ,  $p < 0.001$ ). **B**, Distribution of differences in bandwidth (bw) shown as cell counts.  $iSTRF_{bw}$  was not significantly broader than  $eSTRF_{bw}$  (Kruskal–Wallis test,  $p > 0.05$ ). **C**, Distribution of differences in temporal width (tw) shown as cell counts.  $iSTRF_{tw}$  was significantly larger than  $eSTRF_{tw}$  (Kruskal–Wallis test,  $p < 0.001$ ). **D**, Distribution of differences in STRF latency shown as cell counts. The latency of inhibitory STRFs was significantly longer than that of excitatory STRFs (Kruskal–Wallis test,  $p < 0.001$ ).

nar, cells in LLDp become tuned to ILD through interaction of contralateral excitatory and ipsilateral inhibitory inputs (Manley et al., 1988; Takahashi and Keller, 1992; Mogdans and Knudsen, 1994; Takahashi et al., 1995). Thus, this method allowed us to compute the STRF of the inhibitory input that LLDp neurons receive from the contralateral LLDp and compare it to the STRF of the excitatory input.

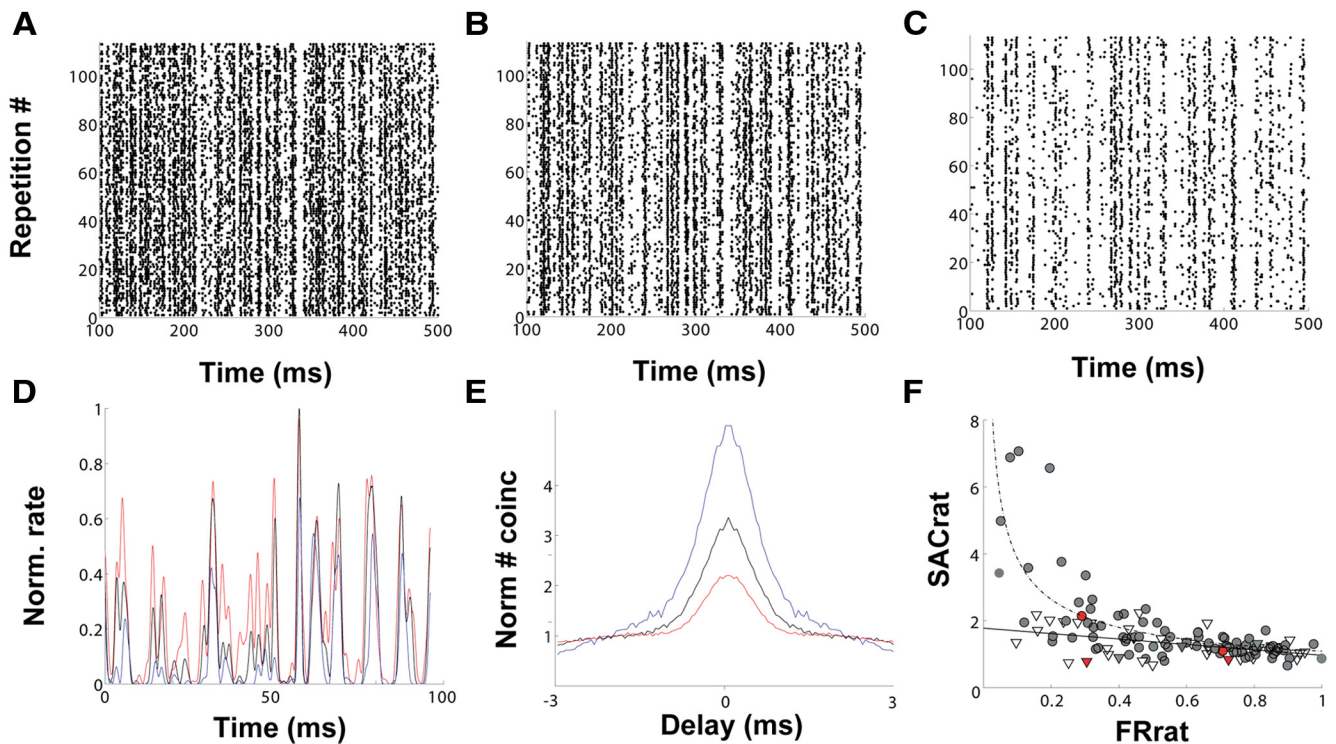
higher in the excitatory ear and comparatively low in the inhibitory ear, leading to a decrease in the latency of excitation and an increase in the latency of inhibition.

#### Response reliability in LLDp

We next studied the effect of inhibition on the response reliability to repeated broadband noise stimuli (frozen noise). Because the inhib-

We fit a GLM of 40 LLDp neurons using their responses to binaurally uncorrelated sounds. The GLM describes the spiking probability as a nonlinear function of a linear combination of the monaural filter outputs, and a spike history term that models refractory processes (Paninski 2004; Calabrese et al. 2011; Plourde et al., 2011). These STRFs were cross-validated with the responses of 34 of these neurons to frozen noise. The mean correlation between the response to frozen noise and the predicted response from monaural STRFs was  $0.58 \pm 0.18$ . The eSTRFs best frequency ( $eSTRF_{bf}$ ) ranged from 1.2 to 8 kHz. The best frequencies of eSTRFs and iSTRFs ( $iSTRF_{bf}$ ) were strongly correlated ( $iSTRF_{bf} = eSTRF_{bf} + 7$  Hz,  $r = 0.96$ ,  $p < 0.001$ , Fig. 4A). The median eSTRF bandwidth was 867 Hz, and the median eSTRF temporal width was 1.47 ms. The median iSTRF bandwidth was 976 Hz; this did not differ significantly from eSTRFs (Kruskal–Wallis test,  $p > 0.05$ ; Fig. 4B). The median iSTRF temporal width was significantly broader than the eSTRF temporal width at 1.85 ms (Kruskal–Wallis test,  $p < 0.001$ ; Fig. 4C). The median latency of the inhibitory minimum, 6.49 ms, was significantly longer than that of the excitatory maximum at 5.47 ms (Kruskal–Wallis test,  $p < 0.001$ ; Fig. 4D).

The difference in latency between excitation and inhibition is consistent with inhibition arriving at LLDp via two synapses (Takahashi and Konishi, 1988; Takahashi and Keller, 1992; Takahashi et al., 1995), while excitation must only traverse one synapse, resulting in a built-in delay. In addition, the latency of excitation and inhibition relative to each other varies with ILD, as the latency of the neural response becomes shorter with increasing sound levels (Aitkin et al., 1970). Thus, as ILD changes, the sound levels at both the excitatory and inhibitory ears change, leading to opposite shifts in latency in both inputs. To measure inhibitory STRFs, ILDs that elicited both excitation and inhibition were used. Most of the data were obtained at small ILDs, at which sound levels at the two ears are similar. Large ILDs favoring the inhibitory side prevent the neuron from firing. At large ILDs favoring excitation, sound levels are much



**Figure 5.** Inhibition enhances response reliability. **A–C**, Representative rasters of one neuron recorded at an ILD that induced minimal (ILD, 20 dB; **A**), intermediate (ILD, 10 dB; **B**), and maximal suppression (ILD, 0 dB; **C**). **D**, Overlapped PSTH excerpts of responses depicted in **A–C** at ILD 20 dB (red), ILD 10 dB (black), and ILD 0 dB (blue). **E**, Normalized SACs for ILD 20 dB (red), ILD 10 dB (black), and ILD 0 dB (blue). **F**, Overlapped data comparing the relationship of normalized SAC peak ratios to corresponding firing rate ratios for changing ILDs (filled circles, dashed line;  $y = 1.09x^{-0.55}$ ,  $r^2 = 0.54$ ,  $n = 51$ ) and monaural stimulation at different sound levels (triangles, solid line;  $y = -0.81x + 1.78$ ,  $r^2 = 0.25$ ,  $n = 38$ ). Data points from the neuron in this figure are highlighted in red. Note that for similar relative decreases in firing rate, proportional changes in response reliability tend to be greater in the presence of inhibition.

itory input to these cells is driven by sound presented to the ipsilateral ear while excitation is driven by sounds in the contralateral ear, the effect of inhibition on response reliability can be assessed by varying sound levels in the inhibitory ear, excitatory ear or both. Changes in firing rate were used as a proxy for the amount of inhibition. Firing-rate changes were normalized across the population by computing changes in firing rate as the ratio of firing rates at different ILDs. We examined the relationship between  $FR_{rat}$  and the corresponding ratio of changes in the normalized SAC peak heights ( $SAC_{rat}$ ).

First, we presented frozen noise at different ILDs. Representative rasters are shown in Figure 5A–C. For each cell, we chose the largest excitatory ILD that elicited the maximum firing rate as the denominator in the  $FR_{rat}$ . Decreasing ILD, i.e., increasing the sound level on the inhibitory input while decreasing the sound level on the excitatory input, causes a decrease in firing rate. ILDs for which the inhibitory input was louder than the excitatory input usually abolished the response. Thus, to measure the effect of inhibition, we presented sound at an ILD in the center of the neuron's dynamic range and an ILD that elicited a lower firing rate but still sufficient to calculate STRFs and SACs. The mean percentage suppression of firing rate was  $45 \pm 25\%$ . We found that response reliability increased as the inhibitory drive became stronger (Fig. 5D,E). There was a power-law relationship between the  $FR_{rat}$  and its corresponding  $SAC_{rat}$  (Fig. 5F;  $y = 1.09x^{-0.55}$ ,  $r^2 = 0.54$ ,  $n = 51$ ).

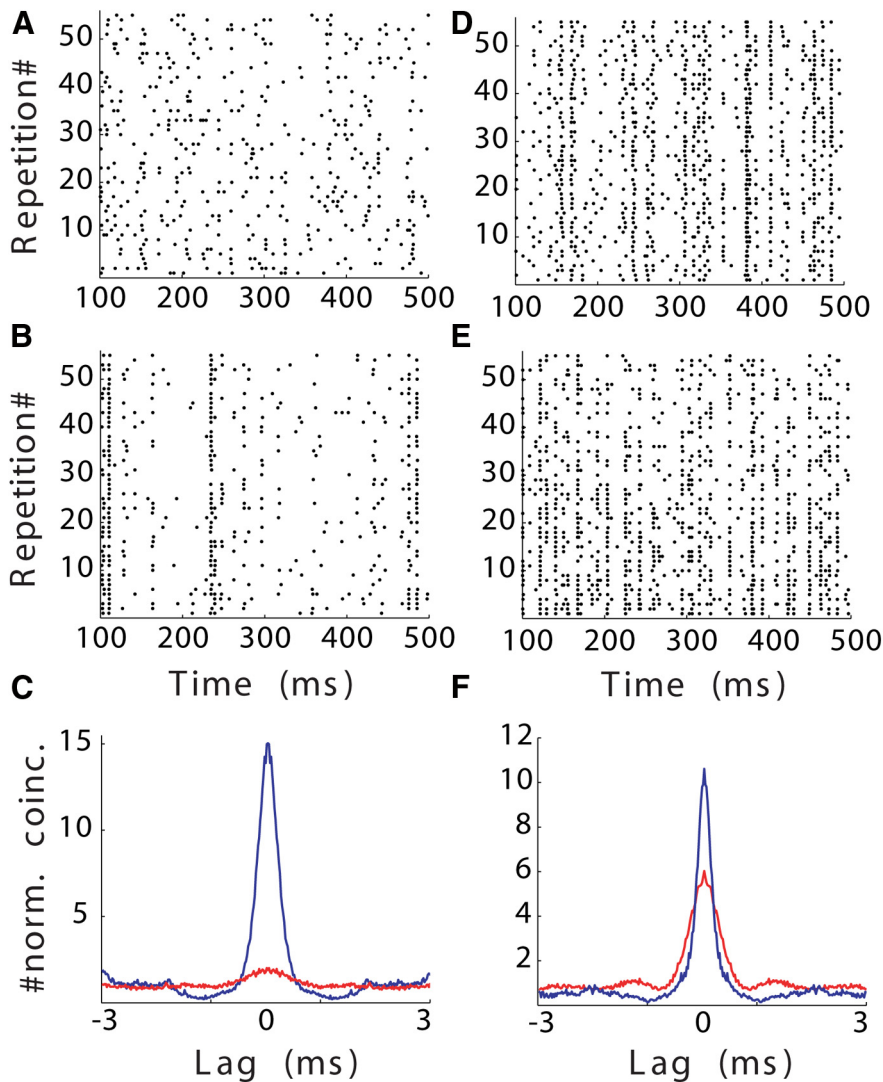
To control for the effects of change in excitatory drive on response reliability, we performed two further experiments. First, we presented frozen noise keeping the sound level in the excitatory ear constant and varying the sound level in the inhibitory ear alone (ILDf curve). The mean percentage suppression of firing rate was  $28.1 \pm 24\%$ . Again we found a power-law relation-

ship between  $FR_{rat}$  and  $SAC_{rat}$  ( $y = 0.92x^{-0.67}$ ,  $r^2 = 0.71$ ,  $n = 25$ ). Next, we presented frozen noise with varying sound levels in the excitatory ear alone with no sound in the inhibitory ear. The mean reduction in firing rate was  $40 \pm 25\%$ . This yielded an inverse relationship between  $FR_{rat}$  and  $SAC_{rat}$  (Fig. 5F;  $y = -0.81x + 1.78$ ,  $r^2 = 0.25$ ,  $n = 38$ ).

To be able to statistically compare these relationships, we converted all data points to a log scale, which allowed us to compare all three stimulus protocols as linear regressions (ILD,  $y = -0.57x + 0.01$ ,  $r^2 = 0.61$ ; ILDf,  $y = -0.58x + 0.06$ ,  $r^2 = 0.45$ ; monaural,  $y = -0.24x + 0.07$ ). We found that both the ILD and ILDf protocol regressions differed significantly from the monaural excitatory protocol regression ( $t$  test,  $p < 0.001$ ). The ILD and ILDf protocol regressions did not differ significantly from each other ( $t$  test,  $p > 0.1$ ).

These analyses were repeated for a subset of neurons from which both monaural excitatory as well as binaural excitatory recordings were available ( $n = 14$ ). Here, we normalized the neurons' firing rate and SAC peak heights by the maximum firing rate observed across all stimulus protocols and its corresponding SAC peak height. The results mirrored those above in that binaural stimulus protocols (ILD, ILDf) yielded a power-law relationship between firing rate and SAC peak height ( $y = 1.1x^{-0.58}$ ,  $r^2 = 0.66$ ), while monaural stimulation yielded a linear relationship ( $y = -0.8x + 1.78$ ,  $r^2 = 0.33$ ). When converted to log scale, the monaural excitatory data yielded a regression described by  $y = -0.25x + 0.06$ , while the binaural protocol data yielded a regression described by  $y = -0.63x - 0.04$ . These regressions are significantly different ( $t$  test,  $p < 0.001$ ).

As a further control, we compared normalized SAC peak heights between binaural and monaural stimulation in a subset of



**Figure 6.** Comparison of monaural and binaural data matched for firing rate for two neurons. **A**, Raster of spikes triggered during monaural presentation of repetitions of the same unique stimulus. Mean FR is 21.4 spikes/s. **B**, Raster of spikes triggered during binaural repetition of a second unique stimulus for the same neuron shown in **A**. Mean FR is 19.5 spikes/s. **C**, Overlapped SACs for monaural (red) and binaural (blue) stimulation protocols for the same cell as in **A** and **B**. **D–F**, Same configuration as in **A–C** for a second neuron. **C**, FR is 39.8 spikes/s. **D**, FR is 40.8 spikes/s. Note that because sound tokens in **A** and **B** and **D** and **E** were different, spiking patterns differ across rasters.

cases where firing rates were within 10% of each other ( $n = 7$ ; Fig. 6). Here we chose conditions that elicited firing rates of 50% or below the maximal firing rate for these neurons to guarantee that in the binaural cases inhibition was present. The 50% value was chosen as an arbitrary cutoff based on the observation that the effect of inhibition on response reliability becomes more prominent at  $FR_{\text{rat}} < 0.5$  (Fig. 5F). The SAC peak height was normalized by the SAC peak observed at each neuron's maximum firing rate across all stimulus protocols ( $SAC_{\text{rat}}$ ). We found that normalized SAC peak heights were significantly larger in the presence of inhibition compared to excitation alone (monaural,  $1.17 \pm 0.56$ ; binaural,  $2.5 \pm 1.73$ ; Kruskal–Wallis test,  $p < 0.001$ ).

These data indicate that inhibition has a significant effect enhancing response reliability, beyond the effect of firing rate modulation induced by changes in excitatory drive alone. This suggests that excitation and inhibition interact in a nonlinear fashion to modulate the gain of spectrotemporal tuning (addressed in the next section).

We compared the SAC measurements to previously published data from the central core of the inferior colliculus (ICcc), the last nucleus belonging exclusively to the ITD-coding pathway (Christianson and Peña, 2007). We found that at ILDs that elicited a maximal firing rate in LLDp, there was no significant difference in response reliability between NA and ICcc (median SAC peak values, LLDp, 3.34; ICcc, 3.74; NA, 4.3; Kruskal–Wallis test,  $p > 0.3$ ). However, for ILDs at which inhibitory drive was present, response reliability in LLDp was significantly greater in LLDp than ICcc (LLDp median SAC peak, 6.11; Kruskal–Wallis test,  $p < 0.03$ ). These data support our hypothesis that the ILD-processing pathway transmits spectrotemporal information with greater fidelity to structures in the midbrain and forebrain than the ITD-processing pathway.

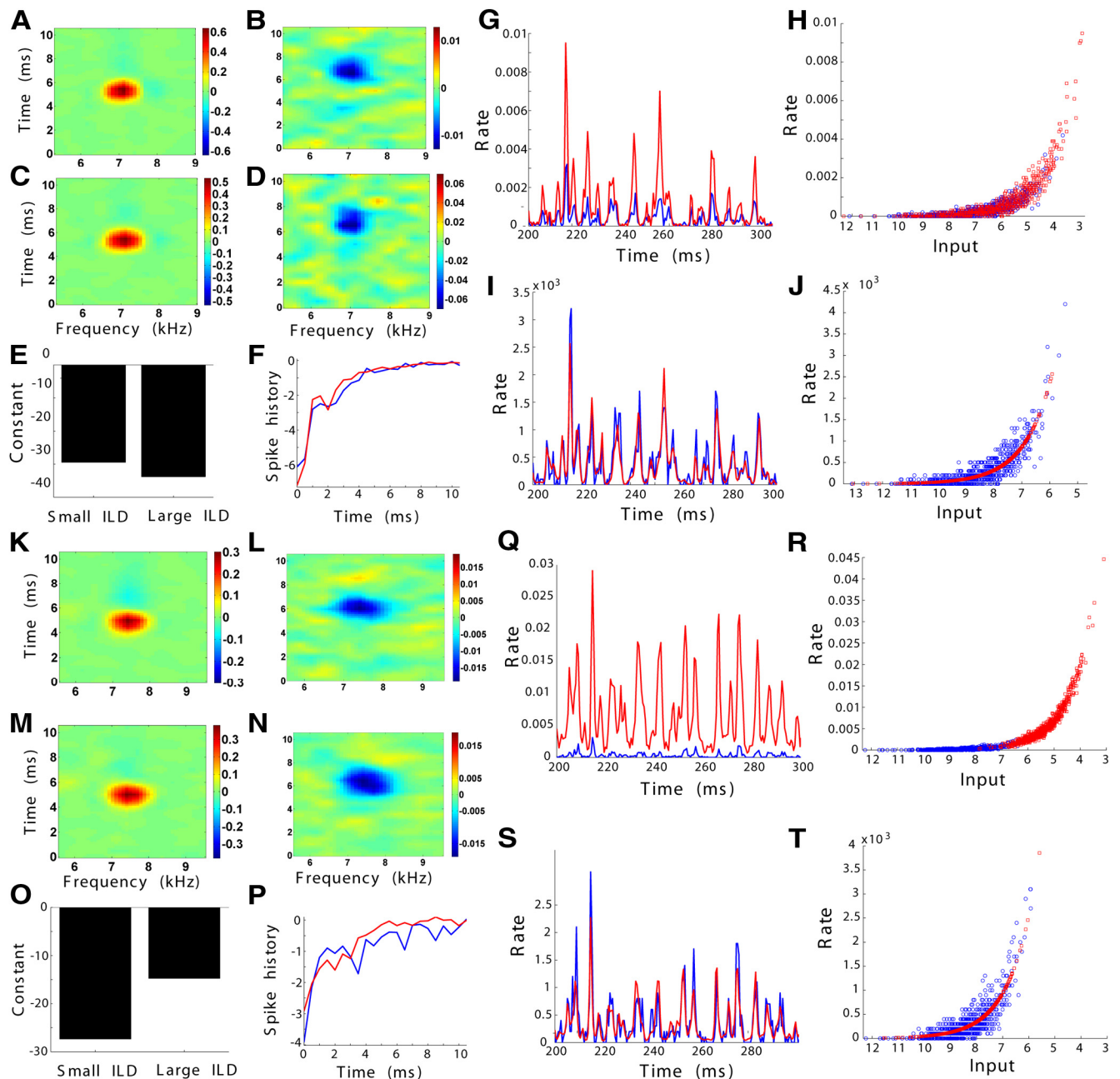
#### ILD-dependent gain modulation

For 19 neurons, we recorded data to fit the GLM at two ILDs (Fig. 7). For those neurons with significant STRFs at both ILDs, the main features of the STRFs including best frequency, bandwidth, latency, temporal width, and peak amplitude did not change significantly when ILD changed ( $n = 15$ ; two-sample Kolmogorov–Smirnov test, all  $p \geq 0.05$ ; Fig. 7A–D, K–N). For four neurons, the inhibitory STRF was not significant at the most excitatory ILD, but emerged at the smaller ILD. The constant bias of the GLM  $k_0$  also did not change significantly with ILD (two-sample Kolmogorov–Smirnov test,  $p = 0.74$ ; Fig. 7E, O). The spike history kernels were also similar across changes in ILD (correlation coefficient,  $0.87 \pm 0.15$ ; Fig. 7F, P). We fit each spike history kernel with an exponential function, and the amplitude and time constant of the exponential fits were not significantly different

between the two ILDs ( $n = 19$ ; two-sample Kolmogorov–Smirnov test,  $p \geq 0.74$  for both). Thus, there was no consistent difference between the component parameters of the GLMs at the two ILDs. However, the overall change in the GLM response with ILD that was consistent across all neurons was a gain modulation.

Increasing inhibition by changing ILD resulted in an overall gain decrease in the firing rate of the neuron. For a frozen noise stimulus, the STRF-filtered input was very similar across ILDs, but there was a change in the gain of the signal (correlation coefficient,  $0.93 \pm 0.07$ ). Correspondingly, the PSTH of the model response were similar across ILDs, but there was a change in gain (correlation coefficient,  $0.78 \pm 0.13$ ; Fig. 7G, I, Q, S). The mean firing rate of the GLM was significantly lower at the more inhibitory ILD (Kruskal–Wallis test,  $p = 0.003$ ).

The GLM responses to frozen noise at the two ILDs were related by a gain modulation. The PSTHs of the GLM responses at the two ILDs showed peaks at the same time points, but the response at the more inhibitory ILD was reduced relative to the



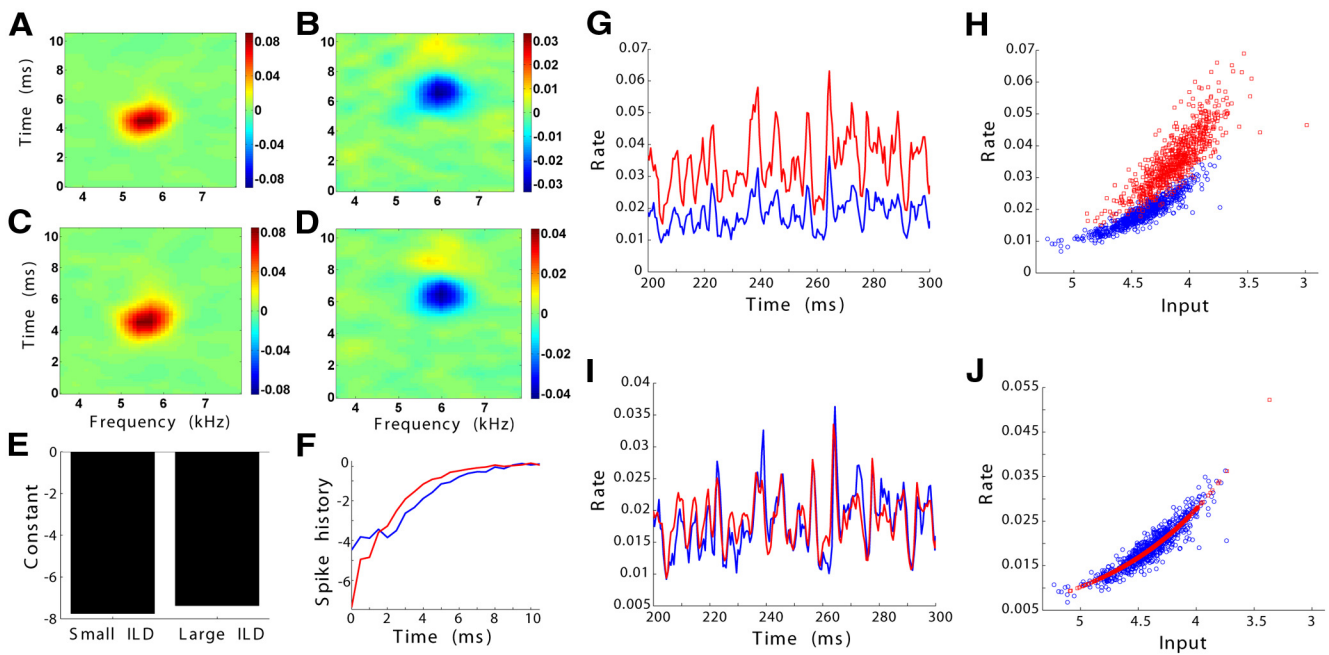
**Figure 7.** Gain modulation of spectrotemporal selectivity as ILD changes. Examples of the generalized linear model at two ILDs for two neurons and the model responses to frozen noise. Here ILD changes by increasing sound level in the inhibitory ear and decreasing sound level in the excitatory ear. **A–F, K–P**, The model components show the eSTRFs (left) and iSTRFs (right) for a large ILD (**A, B, K, L**) and a small, more inhibitory ILD (**C, D, M, N**). The remaining model components are the constant bias (**E, O**) and the spike history kernels (**F, P**). **G–J, Q–T**, The responses of the models to frozen noise on the right reveal an ILD-dependent gain modulation. **G, Q**, The rate responses (spikes per trial in 0.5 ms time bins) of the models at an ILD that induced weaker inhibition (red) and an ILD eliciting stronger inhibition (blue) show peaks at similar times. **H, R**, The input–output functions of the models at the ILDs with (blue circles) and without (red squares) inhibition are similar exponentials, but the inputs lie in different regions. **I, J, S, T**, Applying a gain modulation to the input at the ILD without inhibition (red) predicts the response to the more inhibitory ILD (blue).

response at the other (Fig. 7*G,Q*). The firing rate of the GLM is determined by passing an input signal through an input–output nonlinearity  $f[x(t)]$ , where the input signal consists of the STRF-filtered stimulus and the filtered spike history. Because of the high correlation of the STRF-filtered stimulus across ILDs, it was possible to predict the PSTH at the more inhibitory ILD using a gain-modulated version of the input stimulus at the more excitatory ILD of the form  $g_f[g_2x(t)]$ , where a constant gain is applied to the input signal at the excitatory ILD and to the input–output nonlinearity (Fig. 7*H,J,R,T*; correlation coefficient,  $0.78 \pm 0.12$ ; see Materials and Methods). These data support the conclusion

that spectrotemporal tuning is robust to changes in ILD and that increasing inhibition produces a decrease in neural gain.

To confirm that the gain modulation was an effect of inhibition alone and not of the decrease in excitation induced by changing ILD, we modified the stimulation protocol such that the intensity of the excitatory side was kept constant and only the inhibitory side was changed (ILD*f* curve). As before, model responses in the two conditions were related by a gain modulation. The correlation between the PSTH at the more inhibitory ILD and the rescaled prediction was high (Fig. 8; correlation coefficient,  $0.82 \pm 0.09$ ;  $n = 7$ ).





**Figure 8.** Gain modulation of spectrotemporal selectivity as inhibition increases. Examples of the generalized linear model at two ILDs and the model responses to frozen noise. Here ILD changes by increasing sound level in the inhibitory ear while keeping constant the sound level in the excitatory ear. Plots are as in Figure 7.

Visual inspection of rasters and PSTHs recorded *in vivo* similarly showed that, despite variations in firing rate, cells tended to fire spikes at the same time at different ILDs (Fig. 5D). This supports data from the GLM model that suggest that inhibition does not qualitatively change the spectrotemporal tuning of LLDp cells, but mainly acts to scale responses to spectrotemporal features divisively. To quantify this observation, we calculated the correlation coefficient of the PSTHs for the same stimulus at different ILDs in each cell. The mean correlation coefficient was  $0.54 \pm 0.21$  ( $p < 0.05$ ;  $n = 50$ ) binned at 1 ms. Discrepancies between PSTHs at different ILDs were generally due to decreased spike counts in low-probability bins. This confirms the model's prediction that as inhibitory drive increases and the gain of LLDp neurons decreases, their responses become more selective to spectrotemporal features. Interestingly, the largest peaks in PSTHs did not always correspond across ILDs. Similar effects have been observed in the primary visual cortex under condition of nonclassical receptive field stimulation in conjunction with classical receptive field stimulation, as well as during natural scene presentation (Vinje and Gallant, 2000; Haider et al., 2010).

## Discussion

Here we study how inhibition enhances the reliability of the response to spectrotemporal features of LLDp neurons, the first site of binaural convergence of excitatory and inhibitory inputs in the owl's auditory pathway. These data support previous work in the lateral geniculate nucleus (Butts et al., 2011) as well as previous reports from the auditory cortex (Wehr and Zador, 2003; Wu et al., 2008) and somatosensory cortex (Gabernet et al., 2005; Wilent and Contreras, 2005) that suggest that delayed inhibition can enhance stimulus selectivity, spike-timing precision, and reliability. Our work supports the notion that this is a general mechanism in sensory coding, across stages and species. Our data are unique in that the circuit studied anatomically and functionally separates excitatory and inhibitory inputs to the neurons. This allowed us to clearly delineate the contribution and tuning of the two inputs.

Gain modulation has been studied in a variety of systems, and it has been found that excitation and inhibition act on the input–output function of neurons in a complex manner (Andersen et al., 1985; Chance et al., 2002; Mitchell and Silver, 2003; Murphy and Miller, 2003; Prescott and De Koninck, 2003; Cardin et al., 2008; Ly and Doiron, 2009). The gain modulation effects of inhibition have generally been described in conjunction with noise in the subthreshold membrane potential or the input (Chance et al., 2002; Mitchell and Silver, 2003; Murphy and Miller, 2003; Prescott and De Koninck, 2003; Ly and Doiron, 2009). This is consistent with our data, where the responses of neurons to spectrotemporal features exhibit a fair amount of noise due to both the intrinsic variability of the signal used as stimulus as well as the noisy spike generation. A previous modeling paper by Gittelman et al. (2012) showed that FM selectivity of neurons in the inferior colliculus of the bat may be accounted for by gain modulation via inhibition, which promotes more selective spiking activity. This, too, is consistent with our data, which demonstrate that inhibitory drive exerts divisive gain modulation in LLDp neurons, which decreases spikes fired in low-spiking-probability time bins. Effectively, this leads to greater reliability of these neurons' responses to spectrotemporal features.

Because ILD varies with sound direction, the neurons' gain can be modulated by the spatial location of a stimulus. This mechanism may be useful in auditory scene analysis by adding another dimension to envelope coding (Nelken, 2004). Previous data from Field L, the avian equivalent of auditory cortex, show that the neurons' responses to birdsong may be modulated by spatial location when presented together with a masking sound (Maddox et al., 2012). Interestingly, the influence of spatial location of a sound source was weak in the absence of a masking sound, possibly speaking to the relationship between noise in subthreshold membrane potentials and gain modulation. Our data provide an example of gain modulation at an earlier stage in the auditory processing stream and address how inhibition mechanistically could account for this observation at the level of the brainstem.

The role of stimulus context in gain control has been studied in both the visual and auditory systems. In the visual system, neurons' gain is modulated by contrast (Shapley and Victor, 1981). Andersen et al. (1985) showed that the responses of posterior parietal cortex neurons were gain modulated by gaze angle. Gain modulation is also dependent on the spatiotemporal properties of visual stimuli (Cardin et al., 2008). In the auditory cortex, contrast has been shown to modulate gain (Rabinowitz et al., 2011). Winkowski and Knudsen (2006) showed that top-down gain control modulates responses of neurons in the auditory midbrain, suggesting that attention may play a role.

Another novel aspect of our data set is the relationship between gain modulation and response reliability. A study in the primary visual cortex demonstrated previously that extraclassical receptive field stimulation increases inhibitory postsynaptic potentials, leading to increased reliability of excitatory postsynaptic potentials and spikes (Haider et al., 2010). However, this was not related to gain modulation, as we were able to do in our study. Data from the auditory cortex have shown that combinations of excitation and inhibition with similar spectral tuning could lead to downstream sharpened tuning (Wehr and Zador, 2003; Wu et al., 2008). This may underlie increased spike-timing precision. Our work shows that while spectrotemporal tuning is invariant to the level of excitation and inhibition, inhibition enhances selectivity through gain modulation of the input–output function. This, in turn, leads to enhanced response reliability.

At the cochlear nuclei NM and NA, we found that NA neurons encoded spectrotemporal information with higher fidelity (Steinberg and Peña, 2011). We related this observation to specific attributes of NA cells spectrotemporal tuning, apparent in their STRFs. We found that the greater the magnitude of the suppressive field in NA neurons' STRFs, the greater their response reliability, a measure of their ability to lock to the stimulus envelope. This study examines whether this difference in fidelity of spectrotemporal coding observed between NA and NM is maintained in later processing stages within the sound localization pathways. We show that at the final processing stages of the ILD and ITD pathways, the ILD coding pathway responds to spectrotemporal features with higher fidelity than the ITD pathway.

Mogdans and Knudsen (1994) reported on the discharge pattern of LLDp neurons. Several of their qualitative observations agree with our findings. First, they observed that as sound level in the ipsilateral ear was increased, some spikes during the sustained response disappeared while others remained. Furthermore, spike timing under conditions of strong inhibition was more patterned than when a similar firing rate was elicited by weak excitation only. They showed that while monaural excitatory and binaural stimuli elicited similar firing rates during the onset of the response, the sustained portion of the response exhibited a much lower firing rate during binaural stimulation. This may be due to a temporal delay in the arrival of inhibition at LLDp neurons, thereby shielding the onset of neurons' response. These data support our decision to exclude response onsets from our analysis, by indicating that inhibition most robustly affects the sustained component of the response.

Manley et al. (1988) demonstrated the role of inhibition in the detection of ILD and suggested that excitation and inhibition may be matched in best frequency. It was later shown that the inhibitory drive originated in the contralateral LLDp (Takahashi and Keller, 1992; Takahashi et al., 1995). Takahashi et al. (1995) supported the notion that excitation and inhibition were matched in best frequency, based on the observation that projections from one LLDp targeted the same tonotopic region in the contralateral LLDp. Here we confirm this observa-

tion, as inhibitory and excitatory monaural STRF<sub>bf</sub> values are highly correlated.

LLDp cells project to the lateral shell of the inferior colliculus (ICcl), which also receives input from the ICcc, the last stage of the ITD-processing pathway, and NA. Previous work has demonstrated that cells in ICcl lock well to stimulus envelope with modulation rates of up to 100 Hz (Keller and Takahashi, 2000). Christianson and Peña (2007) demonstrated that neurons in NL, the site of binaural convergence in the ITD-processing pathway, and ICcc also lock to the stimulus envelope, albeit with less reliability than observed in LLDp. It has been suggested previously that the ILD pathway may be responsible for transmitting envelope information to ICcl (Keller and Takahashi, 2000; Steinberg and Peña, 2011). Our data support this hypothesis, as the output of LLDp to ICcl exhibits high-fidelity envelope locking.

It has been suggested that inhibition plays a critical role in encoding and discriminating complex sounds (Narayan et al., 2005; Woolley et al., 2005; Andoni et al., 2007; David et al., 2009; Ye et al., 2010). Inhibition of broader spectral tuning than excitation has been involved in shaping frequency-sweep direction selectivity (Zhang et al., 2003; Kuo and Wu 2012). Moreover, it has been shown that the spectrotemporal profile of the inhibitory drive changes when different types of stimuli are used (Woolley et al., 2005; David et al., 2009). Our work demonstrates that delayed, spectrotemporally tuned inhibition enhances envelope locking by gain control, leading to the suppression of low-probability responses. This is important to the discrimination of complex stimuli, such as speech, whose intelligibility has been shown to rely heavily on envelope information (Shannon et al., 1995; Smith et al., 2002; Fogerty, 2011).

In summary, we demonstrate that neurons in the last stage of the owl's ILD-processing pathway, the nucleus LLDp, encode spectrotemporal features with higher fidelity than the last nucleus of the ITD-processing pathway. Inhibition acts on the LLDp neurons to modulate the gain of the input–output function, which leads to decreased number of spikes in low-probability time bins and increased reliability of envelope locking. Response reliability increases for ILDs where inhibition becomes stronger, yielding a space-dependent gain modulation that enhances envelope locking along the dynamic range of the cells. This gain modulation may constitute an early feedforward mechanism to enhance the reliability of spectrotemporal coding.

## References

- Aitkin LM, Anderson DJ, Brugge JF (1970) Tonotopic organization and discharge characteristics of single neurons in nuclei of the lateral lemniscus of the cat. *J Neurophysiol* 33:421–440. Medline
- Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–458. CrossRef Medline
- Andoni S, Li N, Pollak GD (2007) Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. *J Neurosci* 27:4882–4893. CrossRef Medline
- Arthur BJ (2004) Sensitivity to spectral interaural intensity difference cues in space-specific neurons of the barn owl. *J Comp Physiol A* 190:91–104. CrossRef
- Boudreau JC, Tsuchitani C (1968) Binaural interaction in the cat superior olive S segment. *J Neurophysiol* 31:442–454. Medline
- Butts DA, Weng C, Jin J, Alonso JM, Paninski L (2011) Temporal precision in the visual pathway through the interplay of excitation and stimulus-driven suppression. *J Neurosci* 31:11313–11327. CrossRef Medline
- Calabrese A, Schumacher JW, Schneider DM, Paninski L, Woolley SM (2011) A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. *PLoS One* 6:e16104. CrossRef Medline
- Cardin JA, Palmer LA, Contreras D (2008) Cellular mechanisms underlying stimulus-dependent gain modulation in primary visual cortex neurons *in vivo*. *Neuron* 59:150–160. CrossRef Medline

- Carr CE, Boudreau RE (1991) Central projections of auditory nerve fibers in the barn owl. *J Comp Neurol* 314:306–318. [CrossRef Medline](#)
- Chance FS, Abbott LF, Reyes AD (2002) Gain modulation from background synaptic input. *Neuron* 35:773–782. [CrossRef Medline](#)
- Christianson GB, Peña JL (2007) Preservation of spectrotemporal tuning between the nucleus laminaris and the inferior colliculus of the barn owl. *J Neurophysiol* 97:3544–3553. [CrossRef Medline](#)
- David SV, Mesgarani N, Fritz JB, Shamma SA (2009) Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli. *J Neurosci* 29:3374–3386. [CrossRef Medline](#)
- Escabi MA, Read HL (2003) Representation of spectrotemporal sound information in the ascending auditory pathway. *Biol Cybern* 89:350–362. [CrossRef Medline](#)
- Fischer BJ, Christianson GB, Peña JL (2008) Cross-correlation in the auditory coincidence detectors of owls. *J Neurosci* 28:8107–8115. [CrossRef Medline](#)
- Fischer BJ, Steinberg LJ, Fontaine B, Brette R, Peña JL (2011) Effect of instantaneous frequency glides on interaural time difference processing by auditory coincidence detectors. *Proc Natl Acad Sci U S A* 108:18138–18143. [CrossRef Medline](#)
- Fogerty D (2011) Perceptual weighting of individual and concurrent cues for sentence intelligibility: frequency, envelope, and fine structure. *J Acoust Soc Am* 129:977–988. [CrossRef Medline](#)
- Gabernet L, Jadhav SP, Feldman DE, Carandini M, Scanziani M (2005) Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron* 48:315–327. [CrossRef Medline](#)
- Galambos R, Schwartzkopff J, Rupert A (1959) Microelectrode study of superior olivary nuclei. *Am J Physiol* 197:527–536. [Medline](#)
- Gittelman JX, Wang L, Colburn HS, Pollak GD (2012) Inhibition shapes response selectivity in the inferior colliculus by gain modulation. *Front Neural Circuits* 6:67. [Medline](#)
- Haider B, Krause MR, Duque A, Yu Y, Touryan J, Mazer JA, McCormick DA (2010) Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. *Neuron* 65:107–121. [CrossRef Medline](#)
- Hermes DJ, Aertsen AM, Johannesma PI, Eggermont JJ (1981) Spectrotemporal characteristics of single units in the auditory midbrain of the lightly anaesthetized grass frog (*Rana temporaria* L.) investigated with noise stimuli. *Hear Res* 5:147–178. [CrossRef Medline](#)
- Joris PX (2003) Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* 23:6345–6350. [Medline](#)
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577. [CrossRef Medline](#)
- Joris PX, Louage DH, Cardeon L, van der Heijden M (2006) Correlation index: a new metric to quantify temporal coding. *Hear Res* 216–217:19–30. [Medline](#)
- Keller CH, Takahashi TT (2000) Representation of temporal features of complex sounds by the discharge patterns of neurons in the owl's inferior colliculus. *J Neurophysiol* 84:2638–2650. [Medline](#)
- Konishi M (2003) Coding of auditory space. *Annu Rev Neurosci* 26:31–55. [CrossRef Medline](#)
- Kreeger LJ, Arshed A, MacLeod KM (2012) Intrinsic firing properties in the avian auditory brain stem allow both integration and encoding of temporally modulated noisy inputs *in vitro*. *J Neurophysiol* 108:2794–2809. [CrossRef Medline](#)
- Kuo RI, Wu GK (2012) The generation of direction selectivity in the auditory system. *Neuron* 73:1016–1027. [CrossRef Medline](#)
- Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM (2003) Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* 90:2660–2675. [CrossRef Medline](#)
- Ly C, Doiron B (2009) Divisive gain modulation with dynamic stimuli in integrate-and-fire neurons. *PLoS Comput Biol* 5:e1000365. [CrossRef Medline](#)
- Maddox RK, Billimoria CP, Perrone BP, Shinn-Cunningham BG, Sen K (2012) Competing sound sources reveal spatial effects in cortical processing. *PLoS Biol* 10:e1001319. [CrossRef Medline](#)
- Manley GA, Köppl C, Konishi M (1988) A neural map of interaural intensity differences in the brain stem of the barn owl. *J Neurosci* 8:2665–2676. [Medline](#)
- Mitchell SJ, Silver RA (2003) Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron* 38:433–445. [CrossRef Medline](#)
- Mogdans J, Knudsen EI (1994) Representation of interaural level difference in the VLVp, the first site of binaural comparison in the barn owl's auditory system. *Hear Res* 74:148–164. [CrossRef Medline](#)
- Murphy BK, Miller KD (2003) Multiplicative gain changes are induced by excitation or inhibition alone. *J Neurosci* 23:10040–10051. [Medline](#)
- Narayan R, Ergün A, Sen K (2005) Delayed inhibition in cortical receptive fields and the discrimination of complex stimuli. *J Neurophysiol* 94:2970–2975. [CrossRef Medline](#)
- Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480. [CrossRef Medline](#)
- Nelson BS, Takahashi TT (2010) Spatial hearing in echotic environments: the role of the envelope in owls. *Neuron* 67:643–655. [CrossRef Medline](#)
- Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15:243–262. [CrossRef Medline](#)
- Plourde E, Delgutte B, Brown EN (2011) A point process model for auditory neurons considering both their intrinsic dynamics and the spectrotemporal properties of an extrinsic signal. *IEEE Trans Biomed Eng* 58:1507–1510. [CrossRef Medline](#)
- Prescott SA, De Koninck Y (2003) Gain control of firing rate by shunting inhibition: roles of synaptic noise and dendritic saturation. *Proc Natl Acad Sci U S A* 100:2076–2081. [CrossRef Medline](#)
- Rabinowitz NC, Willmore BD, Schnupp JW, King AJ (2011) Contrast gain control in the auditory cortex. *Neuron* 70:1178–1191. [CrossRef Medline](#)
- Schneider DM, Woolley SM (2010) Discrimination of communication vocalizations by single neurons and groups of neurons in the auditory midbrain. *J Neurophysiol* 103:3248–3265. [CrossRef Medline](#)
- Schnupp JW, Carr CE (2009) On hearing with more than one ear: lessons from evolution. *Nat Neurosci* 12:692–697. [CrossRef Medline](#)
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304. [CrossRef Medline](#)
- Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416:87–90. [CrossRef Medline](#)
- Shapley RM, Victor JD (1981) How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *J Physiol* 318:161–179. [Medline](#)
- Steinberg LJ, Peña JL (2011) Difference in response reliability predicted by spectrotemporal tuning in the cochlear nuclei of barn owls. *J Neurosci* 31:3234–3242. [CrossRef Medline](#)
- Sullivan WE, Konishi M (1984) Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. *J Neurosci* 4:1787–1799. [Medline](#)
- Suta D, Popelár J, Kvasnák E, Syka J (2007) Representation of species-specific vocalizations in the medial geniculate body of the guinea pig. *Exp Brain Res* 183:377–388. [CrossRef Medline](#)
- Takahashi T, Moiseff A, Konishi M (1984) Time and intensity cues are processed independently in the auditory system of the owl. *J Neurosci* 4:1781–1786. [Medline](#)
- Takahashi TT, Keller CH (1992) Commissural connections mediate inhibition for the computation of interaural level difference in the barn owl. *J Comp Physiol A* 170:161–169. [Medline](#)
- Takahashi TT, Konishi M (1988) Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. *J Comp Neurol* 274:190–211. [CrossRef Medline](#)
- Takahashi TT, Wagner H, Konishi M (1989) Role of commissural projections in the representation of bilateral auditory space in the barn owl's inferior colliculus. *J Comp Neurol* 281:545–554. [CrossRef Medline](#)
- Takahashi TT, Barberini CL, Keller CH (1995) An anatomical substrate for the inhibitory gradient in the VLVp of the owl. *J Comp Neurol* 358:294–304. [CrossRef Medline](#)
- Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20:2315–2331. [Medline](#)
- Vinje WE, Gallant JL (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287:1273–1276. [CrossRef Medline](#)
- Wang S, Xu L, Mannell R (2011) Relative contributions of temporal envelope and fine structure cues to lexical tone recognition in hearing-impaired listeners. *J Assoc Res Otolaryngol* 12:783–794. [CrossRef Medline](#)
- Wehr M, Zador AM (2003) Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426:442–446. [CrossRef Medline](#)

- Wilentz WB, Contreras D (2005) Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nat Neurosci* 8:1364–1370. [CrossRef Medline](#)
- Winkowski DE, Knudsen EI (2006) Top-down gain control of the auditory space map by gaze control circuitry in the barn owl. *Nature* 439:336–339. [CrossRef Medline](#)
- Woolley SM, Fremouw TE, Hsu A, Theunissen FE (2005) Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat Neurosci* 8:1371–1379. [CrossRef Medline](#)
- Wu GK, Arbuckle R, Liu BH, Tao HW, Zhang LI (2008) Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. *Neuron* 58:132–143. [CrossRef Medline](#)
- Ye CQ, Poo MM, Dan Y, Zhang XH (2010) Synaptic mechanisms of direction selectivity in primary auditory cortex. *J Neurosci* 30:1861–1868. [CrossRef Medline](#)
- Zar JH (2009) *Biostatistical analysis*. Upper Saddle River, NJ: Pearson Prentice-Hall.
- Zhang LI, Tan AY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424:201–205. [CrossRef Medline](#)