

# Emergence of Spatial Stream Segregation in the Ascending Auditory Pathway

 Justin D. Yao,<sup>1,3</sup>  Peter Bremen,<sup>2,3</sup> and  John C. Middlebrooks<sup>1,2,3,4,5</sup>

Departments of <sup>1</sup>Neurobiology and Behavior, <sup>2</sup>Otolaryngology, and <sup>4</sup>Cognitive Sciences, and <sup>5</sup>Biomedical Engineering, and <sup>3</sup>Center for Hearing Research, University of California–Irvine, Irvine, California 92697

Stream segregation enables a listener to disentangle multiple competing sequences of sounds. A recent study from our laboratory demonstrated that cortical neurons in anesthetized cats exhibit spatial stream segregation (SSS) by synchronizing preferentially to one of two sequences of noise bursts that alternate between two source locations. Here, we examine the emergence of SSS along the ascending auditory pathway. Extracellular recordings were made in anesthetized rats from the inferior colliculus (IC), the nucleus of the brachium of the IC (BIN), the medial geniculate body (MGB), and the primary auditory cortex (A1). Stimuli consisted of interleaved sequences of broadband noise bursts that alternated between two source locations. At stimulus presentation rates of 5 and 10 bursts per second, at which human listeners report robust SSS, neural SSS is weak in the central nucleus of the IC (ICC), it appears in the nucleus of the brachium of the IC (BIN) and in approximately two-thirds of neurons in the ventral MGB (MGBv), and is prominent throughout A1. The enhancement of SSS at the cortical level reflects both increased spatial sensitivity and increased forward suppression. We demonstrate that forward suppression in A1 does not result from synaptic inhibition at the cortical level. Instead, forward suppression might reflect synaptic depression in the thalamocortical projection. Together, our findings indicate that auditory streams are increasingly segregated along the ascending auditory pathway as distinct mutually synchronized neural populations.

**Key words:** auditory cortex; auditory scene analysis; cocktail party; forward suppression; inferior colliculus; medial geniculate body

## Significance Statement

Listeners are capable of disentangling multiple competing sequences of sounds that originate from distinct sources. This stream segregation is aided by differences in spatial location between the sources. A possible substrate of spatial stream segregation (SSS) has been described in the auditory cortex, but the mechanisms leading to those cortical responses are unknown. Here, we investigated SSS in three levels of the ascending auditory pathway with extracellular unit recordings in anesthetized rats. We found that neural SSS emerges within the ascending auditory pathway as a consequence of sharpening of spatial sensitivity and increasing forward suppression. Our results highlight brainstem mechanisms that culminate in SSS at the level of the auditory cortex.

## Introduction

In the complex auditory scenes experienced in everyday life, listeners can disentangle competing sound sequences from multiple

sources, a phenomenon known as “stream segregation” (Bregman, 1990). There is a robust spatial component to this phenomenon, with as little as 8° spatial separation between competing sound sources resulting in perceptual segregation of streams (Middlebrooks and Onsan, 2012). Moreover, neurons in the auditory cortex of anesthetized cats exhibit a correlate of spatial stream segregation (SSS) by synchronizing preferentially to one of two sequences of noise bursts that alternate in location (Middlebrooks and Bremen, 2013). Here, we examine in rats the emergence of SSS among four levels of the ascending auditory pathway: the central nucleus of the inferior colliculus (ICC), the nucleus of the brachium of the inferior colliculus (BIN), the ventral division of the medial geniculate body (MGBv), and primary auditory cortex (area A1). The rat is a suitable experimental model for these experiments because it shows good spatial acuity in psychophysical tests, at least across the frontal midline (Hefner and Hefner, 1985; Kavanagh and Kelly, 1986; Ito et al.,

Received Aug. 17, 2015; revised Oct. 30, 2015; accepted Nov. 6, 2015.

Author contributions: J.D.Y., P.B., and J.C.M. designed research; J.D.Y. and P.B. performed research; J.D.Y., P.B., and J.C.M. contributed unpublished reagents/analytic tools; J.D.Y. analyzed data; J.D.Y., P.B., and J.C.M. wrote the paper.

The work was supported by National Institute on Deafness and Other Communication Disorders Grants R01 DC000420 to J.C.M. and F31 DC013013 to J.D.Y. We thank Zekiye Onsan, Lauren Javier, and Elizabeth McGuire for technical and administrative assistance; and Dr. Raju Metherate for expert advice regarding pharmacological manipulations.

The authors declare no competing financial interests.

This article is freely available online through the *JNeurosci* Author Open Choice option.

Correspondence should be addressed to Dr. John C. Middlebrooks, Department of Otolaryngology, University of California–Irvine, Medical Sciences E, Room E116, Irvine, CA 92697-5310. E-mail: j.midd@uci.edu.

DOI:10.1523/JNEUROSCI.3116-15.2015

Copyright © 2015 Yao et al.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

1996), because neurons in cortical area A1 show homogeneous patterns of spatial sensitivity (Yao et al., 2013), and because mid-brain, thalamic, and cortical levels of the ascending auditory system are readily accessible for study (Yao et al., 2015).

It is known that both spatial tuning (Yao et al., 2015) and the ability of neurons to synchronize to sound envelopes (Creutzfeldt et al., 1980; Joris et al., 2004; Wang et al., 2008) undergo pronounced transformations along the ascending auditory pathway. In the rat, for example, spatial tuning of neurons sharpens and becomes increasingly level-tolerant at successive levels of the pathway from the ICC to a subpopulation of neurons located in the MGBv to essentially all neurons studied in cortical area A1 (Yao et al., 2015). Moreover, the maximum frequencies at which neurons synchronize to stimulus envelopes decrease at successive levels of the pathway, from up to 340 Hz in the IC (rat: Rees and Møller, 1987), to ~100 Hz in MGB (guinea pig: Creutzfeldt et al., 1980; marmoset: Bartlett and Wang, 2007) to ≤18 Hz in cortical neurons (rat: Gaese and Ostwald, 1995; Kilgard and Merzenich, 1999).

We hypothesized that sharpening of spatial sensitivity and decreases in the maximum frequency for envelope synchrony result in enhanced segregation of sequences of sounds from multiple sources, thereby rendering the individual sound streams accessible for subsequent perceptual selection. Accordingly, we tested the hypothesis that SSS strengthens along the ascending auditory pathway due to a gradual increase in spatial sensitivity and a decrease in a low-pass envelope filter cutoff between area A1 and its subcortical input.

We found that, at temporal scales at which spatial stream segregation is seen behaviorally (Middlebrooks and Onsan, 2012), neurons in the ICC showed little evidence of SSS, but that SSS emerges in the BIN and in approximately two-thirds of neurons in the MGBv, and is prominent among all neurons in A1. The SSS observed among neurons in the MGBv and BIN could be explained largely by the sharp spatial tuning of those neurons, whereas the SSS observed in cortical area A1 also reflected a low-pass envelope filter resulting from forward suppression. To further elucidate the mechanism of cortical SSS, we tested the hypothesis that forward suppression observed in the cortex results from GABA-ergic inhibition at the level of thalamocortical synapses. Contrary to that hypothesis, topical application of GABA<sub>A</sub> and GABA<sub>B</sub> receptor antagonists in the auditory cortex gave no relief from forward suppression, despite producing an overall increase in spike rate. We conclude that the forward suppression that we observe in the auditory cortex is due to synaptic depression rather than synaptic inhibition or intrinsic biophysical properties of A1 neurons and that forward suppression combined with the sharpened spatial sensitivity seen at cortical levels result in a neural correlate for SSS in the auditory cortex.

## Materials and Methods

The data reported here were obtained from the population of neurons from which data were presented in previous reports (Yao et al., 2013; Yao et al., 2015). An additional population of neurons was studied using pharmacological procedures. The previous reports focused on frequency and spatial sensitivity measured using tonal stimuli and single broadband noise bursts, respectively. The present study focuses on segregation of interleaved sequences of broadband noise bursts from paired sources.

**Animal preparation.** All procedures were performed with the approval of the University of California at Irvine Institutional Animal Care and Use Committee according to the National Institutes of Health guidelines and were similar to those of previous reports from our laboratory (Middlebrooks and Bremen, 2013; Yao et al., 2013; Yao et al., 2015). Data presented here were obtained from 33 adult male Sprague Dawley rats

(median age: 10.7 weeks; Charles River Laboratories) weighing 265–475 g (median weight: 370 g). The IC and MGB were studied in 15 rats, and cortical area A1 was studied in 18 rats. Surgical anesthesia was induced with urethane (1.5 g/kg i.p.) and xylazine (10 mg/kg i.p.) and supplemented as needed to maintain an areflexive state. To reduce the viscosity of bronchial secretions and to prevent brain edema, we administered atropine sulfate (0.1 mg/kg i.p.) and dexamethasone (0.25 mg/kg i.p.), respectively, at the beginning of surgery and every 12 h thereafter. Core body temperature was maintained at ~37°C. Surgery began with a midline scalp incision and the exposure of the underlying skull. We cemented an inverted machine screw to the skull on the midline, rostral to bregma, to serve as a head holder. The skull was opened to access the right auditory cortex, IC, and MGB. Before recordings, the scalp was partially closed and the positions of the pinnae were adjusted to minimize any alteration that the surgical procedure may have caused.

**Experimental apparatus, stimulus generation, and data acquisition.** The animal was positioned in the center of a darkened double-walled, sound-attenuating chamber (Industrial Acoustics; inside dimensions 2.6 × 2.6 × 2.5 m) that was lined with 60-mm-thick absorbent foam (SONEXone). The animal's head was supported by a 10-mm-diameter rod attached to the skull screw. The rod was held by a thin metal frame positioned behind the animal. The area around the head and ears was unobstructed. A circular hoop, 1.2 m in radius, supported 8.4 cm coaxial loudspeakers (Pioneer Electronics) in the horizontal plane aligned with the rat's interaural axis, 1.2 m above the floor. The loudspeakers were spaced at 20° increments from left to right 80° relative to the rat's midline. Left and right loudspeaker locations are given as contralateral (C) and ipsilateral (I), respectively, with respect to the side of the recording sites, which were all in the right hemisphere.

We used Tucker-Davis Technologies System 3 equipment controlled by a personal computer. Custom MATLAB (The MathWorks) scripts controlled the stimulus sequences, acquired the neural waveforms, and provided on-line monitoring of responses at the 16 or 32 recording sites. Sounds were generated with 24-bit precision at a 100 kHz sampling rate. Loudspeakers were calibrated using a precision microphone positioned in the center of the sound chamber at the normal position of the rat's head; the rat was absent during the calibration. Golay codes (Zhou et al., 1992) were used for calibration of broadband sounds. The broadband frequency responses of the loudspeakers were flattened and equalized such that for each loudspeaker the SD of the magnitude spectrum across the 0.2 to 25 kHz was <1 dB. The stimulus spectrum was rolled off by 10 dB from 25 to 40 kHz. Tonal stimuli were calibrated with tone bursts from 0.2 to 40 kHz (1/6th octave steps).

We recorded extracellular spike activity with single-shank silicon-substrate multisite recording probes from NeuroNexus Technologies using high-impedance head stages and multichannel amplifiers from Tucker-Davis Technologies. The probes had either 16 recording sites spaced at 100 μm intervals or 32 sites spaced at 50 μm intervals; recording-site areas were 413 and 177 μm<sup>2</sup> for 16 and 32 site probes, respectively. Neural waveforms were digitized with 16-bit precision at 25 kHz filtered, and stored on computer disk for off-line analysis.

**Experimental procedure.** Extracellular recordings were performed in area A1 with 16-channel probes, and in the MGB and the IC with 16- and 32-channel probes. Recording procedures are identical to those of previous studies from our laboratory (Yao et al., 2013; Yao et al., 2015). For recordings in A1 (18 rats), the probe was aligned to be approximately orthogonal to the cortical surface and was adjusted in depth to maximize the number of recording sites in active cortical layers. Neural spike activity was encountered along probe segments spanning ~1000 μm in length (median; 5th and 95th percentiles: 500 and 1400 μm). Because the thickness of the rat auditory cortex has been reported to average 1100 μm (Games and Winer, 1988), that means that our cortical recordings consistently sampled the thalamocortical-recipient (granular) layers as well as substantial portions of infragranular and supragranular layers. Cortical area A1 was identified by brisk short-latency responses to noise bursts (first-spike latencies ~10–15 ms), V-shaped frequency tuning curves, and a caudal-to-rostral increase in characteristic frequencies (CFs) (Polley et al., 2007). The borders of A1 were defined by reversals in tonotopy and increases in latencies (Doron et al., 2002; Rutkowski et al., 2003). Our

previous findings demonstrated uniform contralateral hemifield spatial tuning in this neural sample (Yao et al., 2013). Higgins et al. (2010) showed sensitivity to interaural level differences in A1 consistent with contralateral spatial tuning, whereas they found interaural level difference sensitivity consistent with tuning to the spatial midline in neighboring areas VAF and caudal SRAF. That our neural sample showed consistently contralateral sensitivity uncontaminated by neurons showing midline sensitivity supports the view that our cortical data sample was limited to area A1. After each probe was in position, the cortical surface was covered with a warmed 2% solution of agarose in Ringer's solution, which cooled to a firm gel that reduced brain pulsations and prevented the cortical surface from drying.

The right IC was accessed with vertical probe placements (14 placements in 5 rats) ~2–3 mm lateral to the midline and ~7–9 mm caudal to bregma. Two approaches were used to access the right MGB. The vertical approach (22 probe placements in 8 rats) used vertical probe placements ~3–4 mm lateral to the midline and ~5–6 mm caudal to bregma. Four rats used in the vertical approach to the MGB were also used to access the IC. The lateral approach (18 probe placements in 6 rats) used a dorsolateral to ventromedial trajectory, ~30°–50° from the sagittal plane, ~4–6 mm caudal to bregma. Unit localization within the ICC, BIN, and MGBv was based on stereotaxic coordinates and physiological criteria and further confirmed histologically (Yao et al., 2015). Recordings from the ICC were characterized by monotonically increasing CFs along the dorsal to ventral trajectories, with ranges of CFs in individual probes spanning 2.2/2.7/4.1 octaves (25th/50th/75th percentiles). Recordings from the MGBv were characterized by a lateral to medial increase in CF. The border with the medial division of the MGB was marked by an increase in frequency bandwidth and a reversal in tonotopy.

Study of responses at each probe position consisted of measurements of frequency response areas, of mean-spike-rate-versus-azimuth functions (RAFs), of excitation thresholds for broadband noise bursts, and of spatial stream segregation. Measurements were based on single-unit and multiunit responses, as defined in Data analysis. Frequency response areas (FRAs) were measured with pure tones, 80 ms in duration with 5 ms raised-cosine onset and offset ramps, at a repetition rate of 1 or 1.25 s<sup>-1</sup> presented from the loudspeaker located 40° contralateral to the right-sided recording site (C40°). Tones varied in frequency from 0.2 to 40 kHz in 1/3 or 1/6 octave steps and in level in 10 dB steps, typically from -10 to 60 or 70 dB SPL with 10 repetitions per frequency-level combination. In cases in which the lowest threshold for tones was obtained at a frequency of ≥32 kHz, the CF was recorded as ≥32 kHz. The reported thresholds for behavioral detection of tones by rats increases by ~60 dB from 32 to 70 kHz (Kelly and Masterton, 1977; Heffner et al., 1994). For that reason, we infer that the true CFs of units labeled ≥32 kHz were no higher than an octave above ≥32 kHz. The range of CFs across the entire sample was 1 to ≥32 kHz (25th/50th/75th percentiles, A1: 6.70/12/≥32; MGBv: 11.2/18/≥32; ICC: 8/12.7/≥32; BIN: 12.7/32/≥32 kHz). We recorded RAFs for 80 ms Gaussian noise bursts across 360° in azimuth in 20° steps at levels ranging from -10 to 70 dB SPL in 10 dB steps (20 repetitions per combination). All stimuli were presented at a repetition rate of 1.00 or 1.25 s<sup>-1</sup>. We have quantified the frequency and spatial tuning characteristics of the neurons reported here in a previous report (Yao et al., 2015). Noise thresholds were measured using 80 ms Gaussian noise bursts at a 1 or 1.25 s<sup>-1</sup> repetition rate from the C40° loudspeaker, varied in level in 5–10 dB steps, with 20 repetitions per level; a silent condition also was included. Noise thresholds along each recording probe were estimated online, and a modal value was selected. Stimulus levels for subsequent measurements were set 40 dB or more above that modal value. Off-line, noise thresholds were measured using a receiver-operating characteristic (ROC) procedure (see Data analysis), and stimulus levels were computed relative to those thresholds. Across 481 unit recordings in ICC, BIN, MGBv, and A1, the distribution of stimulus levels relative to threshold had 5th, 50th, and 95th percentiles of 25.3, 39.7, and 56 dB, respectively. Along the 72 individual probe placements in those structures, thresholds varied by a median of only 16 dB (5th and 95th percentiles: 3.3 and 28 dB).

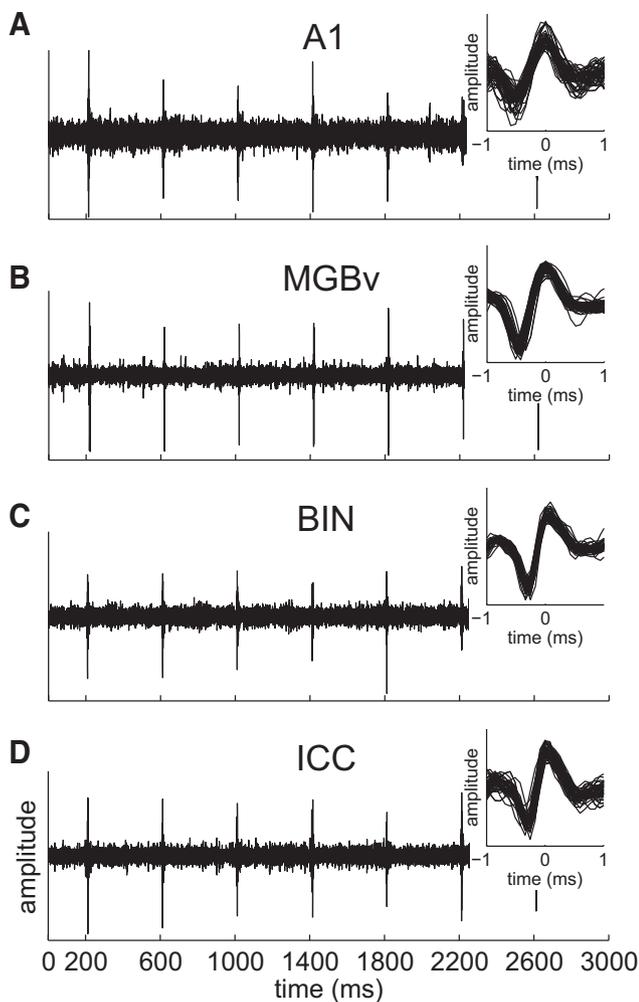
The stimulus conditions for the study of spatial stream segregation were identical to those used in a previous study in our laboratory

(Middlebrooks and Bremen, 2013). In the “competing-source” conditions, stimuli consisted of sequences of independent Gaussian noise bursts, 5 ms in duration with 1 ms cosine-squared on and off ramps. Sequences alternated between A and B sources in an ABAB pattern comprising 15 A and 15 B bursts. Aggregate “base rates” (i.e., presentation rates) of 5, 10, 20, and 40 bursts per second (bps) were tested, such that the difference in onset times between an A burst and that of the following B burst was 200 to 25 ms for base rates of 5 to 40 bps, respectively. The duration of the sequences was 6400 to 800 ms for base rates of 5 to 40 bps, with an additional silent period of ≥700 ms between the offset of one sequence and the onset of the next. The A bursts were presented from C40°, 0°, and ipsilateral 40° (I40°), and the B bursts were presented from C80° to 180° in 20° steps, which included conditions of A and B collocated at C40°, 0°, and I40°. In every case, we also tested B-alone conditions in which the rate of the B stimuli was equal to that in the A-B condition (i.e., half the stated base rate) and in which the rate of B stimuli was equal to that of the aggregate A-B base rate. Every combination of A location (or B alone) and B location was tested once in a random order; then every combination was tested again in a different random order, and so on until every stimulus combination was presented 10 times.

**Pharmacological procedures.** We applied pharmacological agents topically, via a microliter syringe (Hamilton), over the surface of cortical area A1 in 6 of 18 rats to assess the consequences of GABA receptor blockage on responses of A1 neurons. We tested three different types of GABA receptor antagonists: (1) Gabazine (2 rats; 4 probe placements) is a selective postsynaptic GABA<sub>A</sub> receptor antagonist (Heaulme et al., 1986; Kotak et al., 2008); (2) CGP 36216 hydrochloride (2 rats; 4 probe placements) is a selective GABA<sub>B</sub> antagonist that is most active at presynaptic receptors (Ong et al., 2001); and (3) 2-Hydroxysaclofen (2 rats; 6 probe placements) is a selective GABA<sub>B</sub> antagonist that is effective on postsynaptic receptors (Kerr et al., 1988; Metherte and Ashe, 1994). All drugs were freshly dissolved on the day of the experiment. We performed pilot experiments to determine the appropriate drug concentration to be used as well as the time course of their effect on A1 neurons. We found that drug concentrations of ~20–25 μM at a volume of ~4–6 μl could be applied without triggering epileptiform or seizure-like neuronal activity. In these conditions, there was an increase in stimulus-evoked activity starting <10 min after drug application that remained constant for ~45–60 min.

The test of each drug began with placement in the cortex of a multisite recording probe and recording of baseline activity elicited by pulse-train stimuli. Stimuli for the drug tests consisted of sequences of independent Gaussian noise bursts, 5 ms in duration with 1 ms cosine-squared on and off ramps similar to the ABAB patterns used to assess stream segregation. Pulse trains were presented at 12 different repetition rates (1, 2–20, in steps of 2, 30, and 40 bps), 10 times for each rate, with each repetition rate presented once in a random order before repeating all the rates in a different random order. Stimulus levels were set at ≥40 dB above the online estimated modal value from all active channels. After baseline study, one of the GABA antagonists was applied to the cortical surface and the stimulus set was repeated at several after application time points. After study in the drug condition, we washed out the drug with saline and waited ~45 min before shifting the recording probe to another cortical location to test the same or a different drug. In pilot studies, we found no significant difference in spikes per burst between the pre-drug condition compared with the condition ~45 min after drug wash out (Gabazine, paired  $t_{(21)} = 0.69, p = 0.69$ ; CGP 36216, paired  $t_{(26)} = -0.81, p = 0.42$ ; 2-Hydroxysaclofen, paired  $t_{(12)} = -0.90, p = 0.36$ ).

**Data analysis.** All quantitative analyses are based on neuronal action potentials identified with an off-line spike-sorting procedure (for more details, see Yao et al., 2015). Responses were classified as well-isolated single units when they showed the following: (1) consistent waveform appearance upon visual inspection; (2) interspike intervals that revealed a clear refractory period >1 ms; and (3) stability of spike amplitude throughout the recording period. According to that classification, our sample of well-isolated single units consisted of 17 ICC, 13 BIN, 35 MGBv, and 15 A1 neurons. Figure 1 displays neural traces from example single units from A1, MGBv, ICC, and BIN. Those traces were in response to sequences of 5 ms broadband noise bursts presented from C40°



**Figure 1.** Neural traces from example single units from A1 (**A**), MGBv (**B**), BIN (**C**), and ICC (**D**). The traces are responses to sequences of 5 ms broadband noise bursts presented from  $C40^\circ$  at a rate of  $2.5\text{ s}^{-1}$  (i.e., 400 ms ISI) beginning at 200 ms. Insets, Magnified single-spike waveforms (50 random samples) of all responses recorded from those units.

at a rate of  $2.5\text{ s}^{-1}$  (i.e., 400 ms ISI). An additional 401 unit recordings (76 ICC, 52 BIN, 185 MGBv, 88 A1) were classified as multiunit activity consisting of unresolved spikes from two or more neurons. Well-isolated single units are referred to as such and are indicated by symbols in appropriate figures, whereas “unit activity” or “units” refers to single-unit and/or multiunit recordings from a single recording site. All statistics include combined single-unit and multiunit responses, except when stated otherwise. Not all of the stimulus sets were tested for all of the units. Nearly all ICC ( $N = 93$ ), BIN ( $N = 65$ ), and MGBv ( $N = 220$ ) units were tested under all stream segregation base rate conditions, whereas only a subset of A1 units were tested at rates of 20 bps ( $N = 24/103$ ), and no A1 units were tested at 40 bps. In addition to these 481 units, 168 units in A1 were studied using the pharmacological procedures.

We have previously characterized spatial tuning in MGBv (Yao et al., 2015) and reported that neurons in this nucleus show a remarkably bimodal distribution of spatial tuning. Approximately two-thirds of units showed contralateral hemifield spatial tuning, like that seen in A1, and one-third showed omnidirectional spatial tuning like that seen in the ICC. In that study and in the present work, we identified MGBv neurons as either “ICC-like” or “A1-like” using a template-matching procedure based on single-source RAFs (Yao et al., 2015). Briefly, this procedure consisted of comparing the RAF of each MGBv unit with templates of RAFs of ICC and A1 units; all the comparisons used responses to single-burst 80 ms sounds 40 dB above unit thresholds. A similarity index was

computed for each MGBv unit, which indicated whether the MGBv unit was more similar to the ICC (“MGBv-ICC-like,”  $N = 69$ ) or to the A1 (“MGBv-A1-like,”  $N = 151$ ) template.

RAFs for competing-sound stimuli as used in the SSS paradigm expressed mean spike counts per 5 ms noise bursts as a function of loudspeaker location. Spikes tended to fall within  $\sim 25$  ms after each noise-burst onset. For that reason, we counted spikes in the interval 0–25 ms after stimulus onset, which captured essentially all the spikes driven by each A or B noise source sequence. Every stimulus set also included a condition in which the B sequence was presented in isolation either at half, or equal to, the aggregate A-B rate. Mean spikes per burst were computed across 10 trials.

The breadth of spatial sensitivity by each unit was represented by the width of its equivalent rectangular receptive field (ERRF) (Lee and Middlebrooks, 2011; Middlebrooks and Bremen, 2013; Yao et al., 2013; Yao et al., 2015). The ERRF width was computed by integrating the area under a unit’s RAF, forming a rectangle having a peak height and area equal to that of the RAF, and measuring the resulting width.

Procedures based on signal detection theory (Green and Swets, 1966; Macmillan and Creelman, 2005) were used to quantify the discrimination of sound-source location. We accumulated spike counts for all repetitions synchronized to each of the A and B sources and formed an empirical ROC curve based on the trial-by-trial distributions of spike counts elicited on all trials by each of the two stimuli. The area under the ROC curve gave the probability of correct discrimination of the stimuli. That probability was expressed as a  $z$ -score and was multiplied by  $\sqrt{2}$  to obtain the discrimination index,  $d'$ . In some cases, 100% of the spike counts elicited by one stimulus were greater than any of those elicited by the other stimulus, the area under the ROC curve was 1.0, and the corresponding  $z$ -score was undefined. In those cases,  $d'$  was written as  $\pm 2.77$ , corresponding to 97.5% correct discrimination. The sign convention was that  $d'$  was positive when the more contralaterally located sound elicited more spikes.

We used two approaches to quantify spatial stream segregation. In one approach, we measured the difference in spikes per burst synchronized to the A source located at  $C40^\circ$  corresponding to a shift in the B source from colocation at  $C40^\circ$  to spatially separated at  $I40^\circ$ . This was quantified by the Spatial Release Index (SRI), which was as follows:

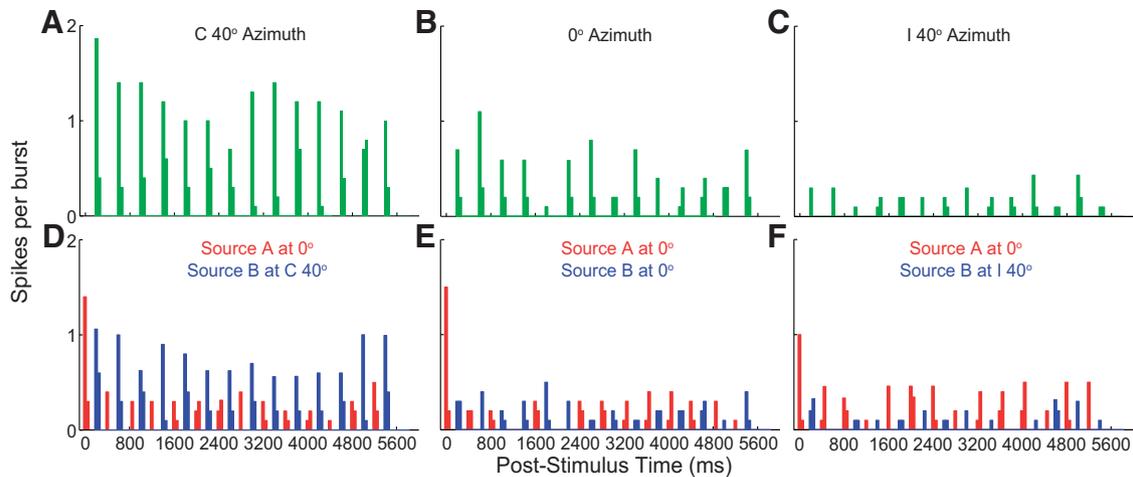
$$\text{SRI} = (R_{I40} - R_{C40}) / (R_{I40} + R_{C40}),$$

where  $R_{I40}$  and  $R_{C40}$  were the responses synchronized to the A stimulus when the B source was at ipsilateral or contralateral  $40^\circ$ , respectively. Positive values of the SRI indicated that separation of A and B sources resulted in a release from masking of the A source.

In the other approach, the magnitude of spatial stream segregation in conditions of interleaved A and B noise bursts was quantified by computing  $d'$  for spikes synchronized to the A versus B bursts. Values of  $d'$  were plotted as a function of B-source location, and source-separation thresholds were taken as the corresponding interpolated separation at which the plot of  $d'$  versus azimuth crossed  $d' = 1$ .

The excitation thresholds (dB SPL) for detection of neural activity elicited by noise bursts were estimated by computing  $d'$  for pairs of noise bursts at successively increasing levels, plotting the cumulative  $d'$  versus sound level and taking as threshold the minimum interpolated (1 dB steps) sound level at which  $d' = 1$ . To obtain each unit’s CF, the matrix of  $d'$  values across all tested frequencies and levels was first screened to eliminate isolated values of  $d' > 1$  for which all neighboring values were  $< 1$ ; this eliminated isolated values lying outside the FRA. The frequency tuning curve (i.e., the border of the FRA) was found by interpolating the  $d'$  values across all tested sound levels in 1 dB steps at each tested frequency and finding the minimum sound level at which  $d'$  was  $\geq 1$ . The CF was given by the frequency of the lowest-level tip of the frequency tuning curve. Again, when an FRA showed its lowest threshold at  $\geq 32$  kHz, the CF was indicated as “ $\geq 32$  kHz.”

The strength of stimulus synchrony of all units for single-source conditions across 2.5 to 40 bps was represented by vector strength (VS; Goldberger and Brown, 1969). The VS could range from 0 (no synchrony)



**Figure 2.** Poststimulus time histogram of a well-isolated single unit from cortical area A1. Plotted are responses to sound sequences from a single-source location (top row, **A–C**) or to competing sequences from two locations (bottom row, **D–F**). Responses are spikes per sound burst in 50 ms bins, averaged over 10 trials. Top row, Single sound-source was located at C40° (**A**), 0° (**B**), I40° (**C**). Bottom row, Red and blue color bars represent spike rates synchronized to the A- and B-source, respectively. In those cases, the A-source was fixed at 0°, whereas the location of the B-source varied from C40° (**D**), 0° (**E**), I40° (**F**). Unit 1216.2.4.

to 1 (all spikes at identical phase). The statistical significance of the VS was evaluated by the Rayleigh test of uniformity (Mardia, 1972) at the level of  $p < 0.001$ .

Statistical procedures used custom-written MATLAB scripts (The MathWorks) that incorporated the MATLAB Statistics Toolbox. *Post hoc* multiple comparisons used the Bonferroni correction. Error bars in the illustrations indicate SD unless stated otherwise.

## Results

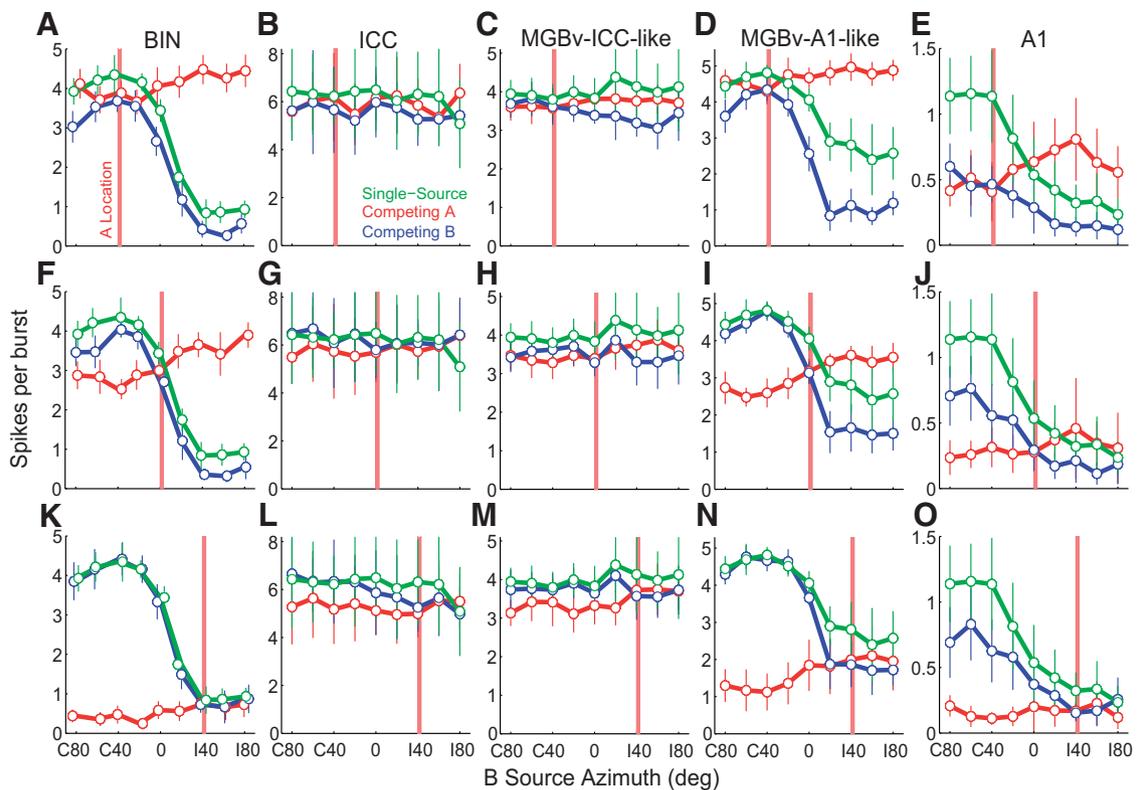
### Neural responses to competing sound sequences

We tested an ABAB stimulus pattern consisting of sequences of noise bursts alternating between two source locations. Poststimulus time histograms from an isolated single unit in cortical area A1 in response to such stimuli are shown in Figure 2. In that example, the “base rate” was 5 bps, referring to the aggregate of A and B rates. Figure 2A–C shows conditions in which B was presented in the absence of A; we refer to this as the “single-source” condition, equivalent to 2.5 bps (half of the aggregate AB base rate of 5 bps). This unit displays contralateral hemifield spatial tuning, with strong responses synchronizing to the single sound source at C40° (Fig. 2A), a decrease in overall spikes per burst to the sound source at 0° (Fig. 2B), and very weak responses to the sound source at I40° (Fig. 2C). Responses to the competing-source condition in which the A source was added at 0° are represented in Figure 2D–F. Under the condition when A and B sources were collocated at 0° (Fig. 2E), the presentation rate was equivalent to a 5 bps single-source condition. The unit responded reliably only to the first sound burst and showed only weak responses to subsequent A (red) and B (blue) bursts. When the B source was shifted to C40° (Fig. 2D), the unit displayed robust responses synchronized to the B source in preference over the A source. Thus, the neural response was captured by the B source, relative to the A source. Under the condition when the B source was shifted to I40° (Fig. 2F), responses to both A and B sources were relatively weak, but responses synchronized to the A source were stronger than those synchronized to the B source.

Synchronized mean spikes per burst from the same A1 single unit in Figure 3 are quantified for the full range of A and B location combinations in the right column of Figure 3. Figure 3E, J, and O shows conditions in which the A sources were fixed at locations C40°, 0°, and I40°, respectively. The green line repli-

cated in all three panels shows the tuning to the location of a single sound source, which exhibited the contralateral hemifield pattern that was characteristic of all our recorded units in A1. This was consistent with the dominant contralateral hemifield tuning to single noise bursts seen in rat A1 units (Higgins et al., 2010; Yao et al., 2013). The location of the B source, plotted on the horizontal axis, influenced both the responses synchronized to the B source (blue lines) as well as those synchronized to the A source (red lines). Moreover, when A and B sources were collocated at the azimuths indicated by the vertical red lines, responses synchronized to A and B were approximately equal, and the numbers of spikes per stimulus burst were reduced to approximately half of that seen for single sources, as indicated by the green lines at the corresponding azimuth. This reduction in response demonstrates forward suppression, in which responses of the A1 unit declined with increasing stimulus presentation rate. As A and B sources were moved apart, the response to one of the sources increased and the other decreased or remained relatively weak. In that way, responses of the A1 unit were captured by one source over the other. The vertical distance between the blue (B-source) and red (A-source) curves at each azimuth of the B source represents the degree to which the sequences of sounds from A and B sources were segregated by the synchronized responses of this neuron.

We examined whether neural responses from subcortical levels displayed SSS similar to that seen in cortical area A1. Responses from one single unit in the ICC (Fig. 3B, G, L) is representative of our population of ICC units in that it displayed little or no spatial sensitivity to single sources. Unlike the case of cortical recordings (e.g., Fig. 3E, J, O), spikes per burst in this ICC unit to the collocated A and B conditions were similar to those seen under single-source conditions presented from the same location. That is, there was little reduction in responses seen between the 2.5 bps single-source and 5 bps competing-source conditions, which demonstrates an absence of forward suppression. Also, the location of the B source had relatively little influence on the responses to the A source such that synchronized responses to the competing sources remained fairly equivalent across spatial configurations. Figure 3 also displays responses from one single unit in the BIN (Fig. 3A, F, K), which lies in the tectal pathway



**Figure 3.** Responses to single-source and competing-source conditions. Mean spike rates ( $\pm 1$  SD) are plotted as a function of B-source location from example single units from BIN (**A, F, K**; unit 1402.4.18), ICC (**B, G, L**; unit 1403.5.31), MGBv-ICC-like (**C, H, M**; unit 1312.2.15), MGBv-A1-like (**D, I, N**; unit 1311.2.13), and cortical area A1 (**E, J, O**; same unit as Fig. 1). Green lines indicate spike rates synchronized to single-source sound sequences with a base rate of 2.5 bps. Red and blue lines indicate responses synchronized to A- and B-sound sources, respectively, in competing-source conditions; the base rate was 5 bps. B-source locations were as plotted, whereas A sources were fixed at the locations indicated by the vertical red lines (**A–E**, C40°; **F–J**, 0°; **K–O**, 140°). Variation in A responses reflected the effect of changing B-source locations.

from the ICC to the superior colliculus. Spatial sensitivity for this particular unit and all BIN units showed sharp contralateral hemifield tuning under the single-source condition. As in the ICC, spikes per burst to the collocated A and B competing sources resembled those seen under single-source conditions presented from the same locations. However, responses under the competing A and B condition preferentially synchronized to either the A or B source, depending on the spatial configuration. With the A source located at 140°, responses synchronized to the B source were indistinguishable from single-source responses. Accordingly, the A source did not influence the neuronal response to the B source.

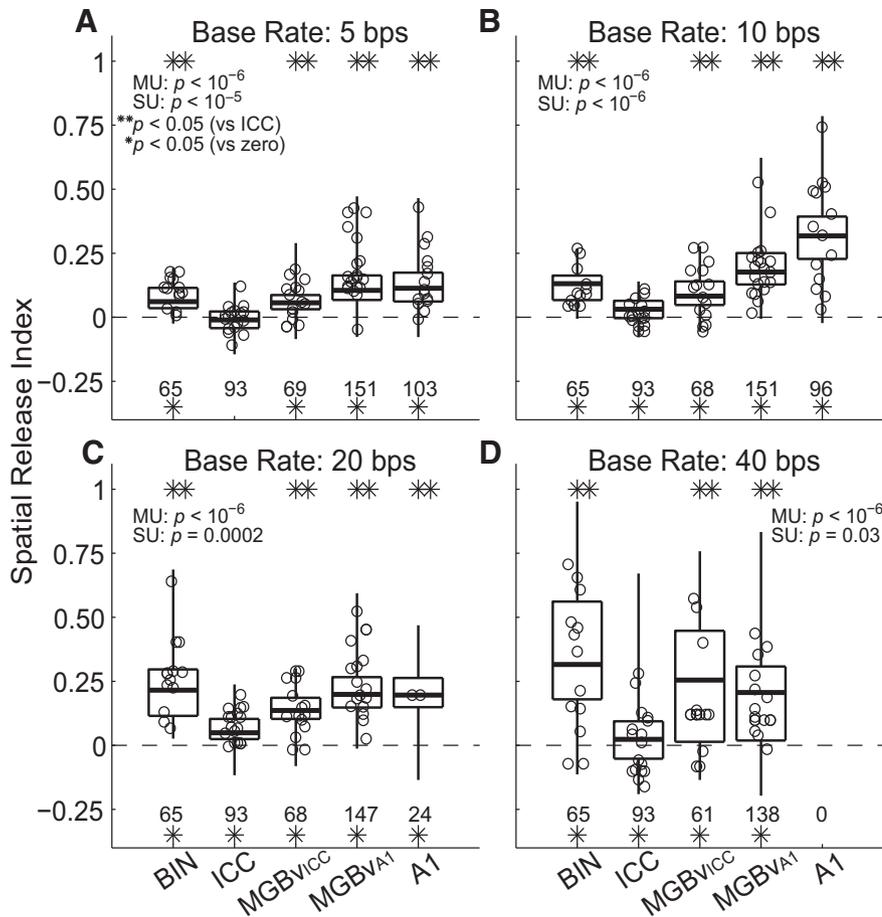
We previously encountered two subpopulations of units in the MGBv, with one subpopulation showing spatial sensitivity similar to that seen in A1 and another one with spatial tuning characteristics similar to ICC (Yao et al., 2015). Those unit populations were denoted as “MGBv-A1-like” and “MGBv-ICC-like” and were distinguished by a quantitative procedure described in Materials and Methods; in the present study, approximately one-third of units were classified as MGBv-ICC-like and two-thirds were MGBv-A1-like. The differences in spatial sensitivity seen between the two subpopulations of units can be seen in their responses to the single source sound sequences in Figure 3. The MGBv-A1-like unit (Fig. 3*D, I, N*) displays contralateral hemifield tuning similar to the single source condition seen in the A1 unit (Fig. 3*E, J, O*), whereas the MGBv-ICC-like unit (Fig. 3*C, H, M*) displays omnidirectional spatial tuning similar to that seen in the ICC unit (Fig. 3*B, G, L*). Overall, we found that responses from the MGBv-A1-like unit showed some SSS by

preferentially synchronizing to either the A or B competing source, whereas responses to competing A and B sources from the MGBv-ICC-like unit were undifferentiated. This difference in SSS is consistent with their differences in spatial sensitivity. Similar to the ICC and BIN, both MGBv units displayed little evidence of forward suppression at the illustrated stimulus base rates, meaning that there was little difference in spike count between collocated competing A and B sources versus the single-source sequence presented from the same location. That is, doubling the stimulus rate by adding sounds from a collocated source had little effect on the rate of spikes per sound burst.

#### Quantification of SSS

One measure of the magnitude of SSS is the effect of B source location on the response to a fixed A source. Specifically, we measured the difference in spikes per bursts synchronized to the A source located at C40° corresponding to a shift in the B source from collocated at C40° to 140°. This was quantified by the SRI (see Materials and Methods). Positive values of the SRI indicated that separation of A and B sources resulted in a release from masking of the A source. The magnitudes of the SRI tended to vary with stimulus presentation rate. All ICC units and nearly all MGB units were tested at base rates of 5, 10, 20, and 40 bps. No A1 units responded to the fastest rate, so A1 units were not routinely tested at 40 bps and few were tested at 20 bps.

Distributions of SRI from all tested units in each population are shown in Figure 4 across base rates of 5, 10, 20, and 40 bps. Generally, spatial release tended to be high among A1 units and lowest among ICC units. At each base rate, the distributions of



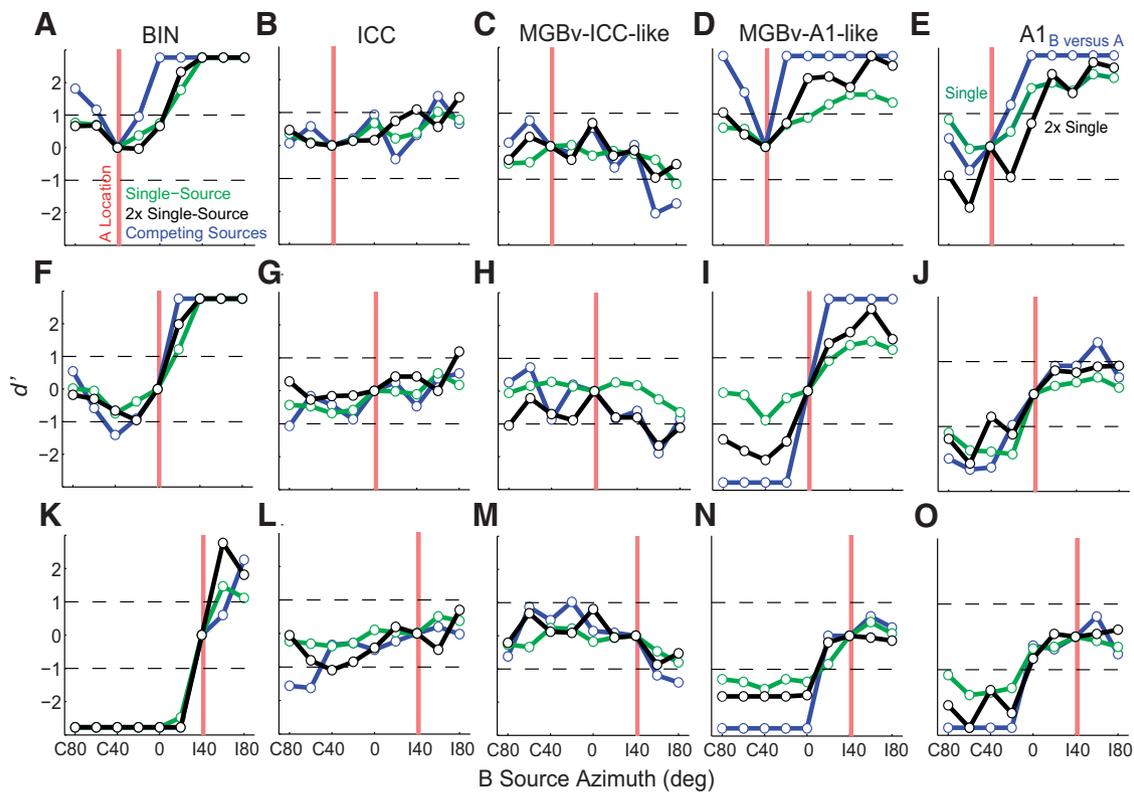
**Figure 4.** Distribution of the SRI values across each unit population for each base rate tested (A, 5 bps; B, 10 bps; C, 20 bps; D, 40 bps). Horizontal lines forming the boxes indicate the 25th, 50th, and 75th percentiles of multiunit responses. Vertical lines indicate the full range from multiunit responses. Circles represent data point from well-isolated single units. *p* values in each subplot indicate results of a nonparametric ANOVA (Kruskal–Wallis) test across the unit populations. MU and SU indicate statistics from multiunit and single-unit recordings. \*\*Significant difference in distribution versus ICC (Bonferroni-corrected multiple comparisons). \*Significant difference for a mean different from 0 (*t* test, Bonferroni-corrected multiple comparisons). The listed numbers under each unit distribution correspond to the number of units in each population.

SRI varied significantly across unit populations (MU:  $\chi^2 = 138–207.4$ ,  $p < 10^{-6}$ ; SU:  $\chi^2 = 24.4–35.9$ ,  $p < 10^{-5}$ ; Kruskal–Wallis) with ICC units displaying the lowest SRIs ( $p < 0.005$ ; Bonferroni-corrected). In A1, SRIs were significantly highest at 10 bps ( $p < 0.0001$ , Bonferroni-corrected), whereas at 5 and 20 bps, SRIs of A1 units were not significantly different from those of BIN and MGBv-A1-like units ( $p > 0.05$ , Bonferroni-corrected); A1 units were not tested at 40 bps. SRIs were higher among MGBv-A1-like units than among MGBv-ICC-like units at 5 and 10 bps ( $p < 0.0001$ , Bonferroni-corrected), whereas there was no significant difference between those populations at 20 and 40 bps. Units in A1 typically showed greater spatial release with increasing stimulus rates. For example, the SRIs more than doubled between base rates of 5 and 10 bps; SRIs in A1 declined again at 20 bps, which might reflect the generally poorer responses of A1 neurons to high-rate stimuli. Subcortical units in the BIN and MGBv also displayed progressively greater spatial release across increasing stimulus rates (SRI medians across 5–40 bps rates: BIN = 0.08, 0.13, 0.23, 0.30; MGBv-A1-like: 0.11, 0.18, 0.20, 0.19; MGBv-ICC-like: 0.05, 0.08, 0.14, 0.26). Median SRI values for ICC units were  $\sim 0$  across all bps conditions, although a small proportion of ICC units (10 of 93; 8 MUs, 2 SUs) at the 40 bps condition exhibited a high degree of spatial release (SRI > 0.20)

that was similar to the other unit populations. These trends indicate a strong relationship between presentation rate and the magnitude of SSS, possibly reflecting the timescale of forward suppression. We examine forward suppression in a later section.

We quantified the discrimination between sound source locations with a discrimination index,  $d'$ . Specifically,  $d'$  quantified the acuity with which sound sequences from the A versus B sources could be discriminated on the basis of trial-by-trial spike rates synchronized to the A or B source (Fig. 5; same example units as in Fig. 3). The blue lines indicate discrimination between A and B competing sources at a given A-source location (indicated by the vertical red line). The green and black lines indicate single-source conditions compared with the three fixed A-source locations (C40°, Fig. 5A–E; 0°, Fig. 5F–J; I40°, Fig. 5K–O); black and green lines indicate stimulus rates equal to the aggregate A–B rate (5 bps) or half that rate (2.5 bps), respectively. Similar to the trends in the responses seen in Figure 3, significant source segregation indicated by magnitudes of  $d' \geq 1$  was typically seen for most non-zero A and B source separations among BIN, MGBv-A1-like, and A1 units. Specifically, significant source segregation was achieved at the minimum A and B source separations that were tested (i.e., 20° separation) when the A-source was fixed at C40° (Fig. 5A, D, E) and 0° (Fig. 5F, I, J). In essentially all conditions, the  $d'$  for spatial segregation was greater for the competing-source (blue line) than for the single-source condition at either base rate (green and black lines). Significant source segregation among ICC (Fig. 5B, G, L) and MGBv-ICC-like (Fig. 5C, H, M) units was seen only at the extreme source separations ( $\sim 80^\circ$  separation). The trends seen for these example units are quantified below for the population.

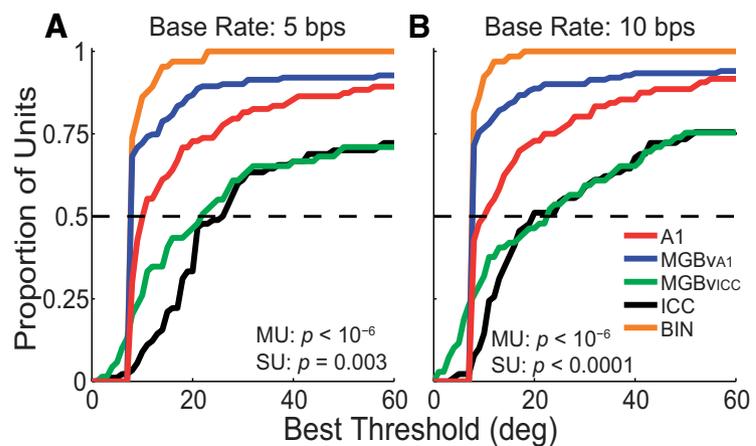
We used the source separation at which  $d'$  crossed  $\pm 1$  (dashed lines) as the spatial threshold for significant segregation between competing A and B sources; values were interpolated in 1° steps. We then selected for each unit the minimum threshold across conditions of A source at C40°, 0°, and I40°; distributions of those minimum thresholds across all unit populations are shown in Figure 6. We found a significant difference in thresholds across all unit populations at 5 bps (Fig. 6A) and 10 bps (Fig. 6B) conditions (MU:  $\chi^2 = 108–130$ ,  $p < 10^{-6}$ ; SU:  $\chi^2 = 25–26$ ,  $p < 0.0001$ ; Kruskal–Wallis). Generally, thresholds were narrowest (i.e., highest-acuity segregation) for BIN and MGBv-A1 units, intermediate for A1 units, and broadest (i.e., worst) for MGBv-ICC and ICC units. *Post hoc* multiple comparisons indicated that, at 5 and 10 bps, thresholds were narrower for BIN and MGBv-A1 units than for those of A1 units ( $p < 0.05$ ; Bonferroni-corrected), and that thresholds of A1 units were narrower than those of MGBv-ICC and ICC units ( $p < 0.05$ ; Bonferroni-corrected). Median values of the distribution of thresholds were very similar among A1, MGBv-A1-like, and BIN units,



**Figure 5.** Discriminating competing sources.  $d'$  values are plotted as a function of B-source location from example single units from BIN (**A, F, K**), ICC (**B, G, L**), MGBv-ICC-like (**C, H, M**), MGBv-A1-like (**D, I, N**), and cortical area A1 (**E, J, O**); same example units plotted in Figure 2. Blue lines indicate  $d'$  values for discrimination of trial-by-trial spike counts elicited by A versus B sources. Green and black lines indicate discrimination of spikes per burst elicited from the plotted location versus those elicited in separate trials from the fixed location indicated by the vertical red line at 2.5 and 5 bps single-source rates, respectively. Similar convention and same units as in Figure 3.

whereas the cumulative distributions diverge among the units having broader minimum thresholds.

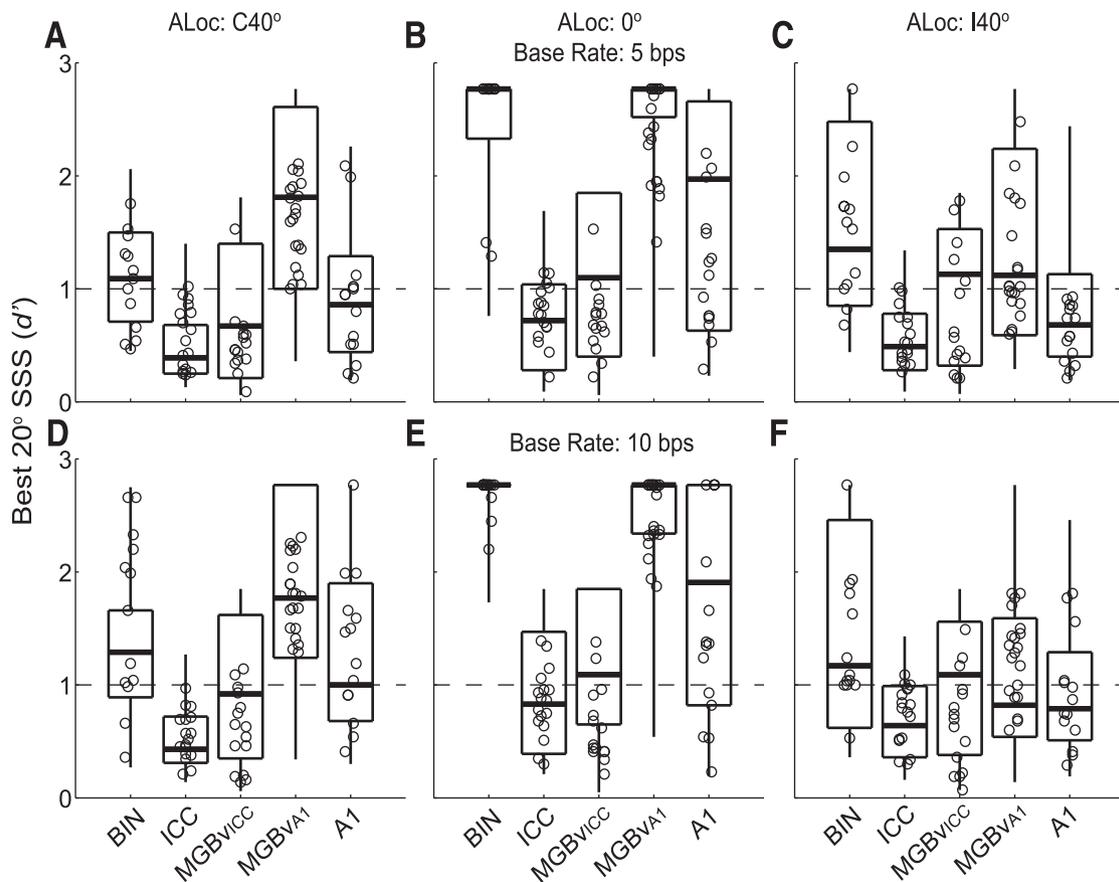
Spatial segregation for various A-source locations was quantified by computing the  $d'$  for discrimination of A or B sources that were separated by 20°. In Figure 7, each box displays the distribution of  $d'$  for discrimination of A and B sources across every combination of A-source location at C40° (Fig. 7A,D), 0° (Fig. 7B,E), or 140° (Fig. 7C,F) and base rate (5 bps: Fig. 7A–C; 10 bps: Fig. 7D–F). For each unit at each A location, the B location resulting in the greater magnitude of  $d'$  for a B location 20° to the left or right of the A source was selected and represented in the distribution by its absolute value. Similar to the spatial release (SRI) results (Fig. 4), A1, MGBv-A1-like, and BIN units showed the greatest segregation between competing A and B sources across all conditions. In addition, spatial stream segregation was strongest when the A-source location was fixed on the midline (Fig. 7B,E). For midline A-source locations, >70% of A1, >85% of MGBv-A1-like units, and >94% of BIN units showed significant spatial stream segregation ( $d' \geq 1$ ), whereas only slightly more than half of MGBv-ICC-like units (54%) and approximately one-third (34%) of ICC units showed significant spatial stream segregation. Overall, these findings indicate that spa-



**Figure 6.** Cumulative distributions of best thresholds (minimum interpolated separation (°) at which  $d'$  crosses  $\pm 1$  across the three conditions of A sources at C40°, 0°, and 140°) for segregating competing A and B sources for each unit population (color coded) at 5 (**A**) and 10 (**B**) base rates. Horizontal dashed line indicates the 50th percentile. Statistics from a nonparametric ANOVA (Kruskal–Wallis) test across the multiunit and single-unit populations are shown in each subplot.

tial stream segregation was strongest among A1, MGBv-A1-like, and BIN units and weakest for MGBv-ICC-like and ICC units.

The tendency of MGBv-A1-like units to show stronger and higher-acuity SSS than units in A1 must reflect to some degree the procedure by which they were selected. That is, the MGBv-A1-like units were a subpopulation of MGBv units selected for similarity to the average of A1 responses, which showed hemifield tuning. The A1 population, in contrast, included the entire A1



**Figure 7.** Distribution of  $d'$  for discrimination of A and B sources. Each  $d'$  in the distributions represents for one unit the greater absolute value of  $d'$  for discrimination of A and B sources across locations 20° to the left or right of A sources at C40° (A, D), 0° (B, E), and 140° (C, F). Rows represent 5 bps (A–C) and 10 bps (D–F) base rates. Each subplot represents boxplot distributions for each unit population. Horizontal lines forming the boxes indicate the 25th, 50th, and 75th percentiles of multiunit responses. Vertical lines indicate the full range of multiunit responses. Circles represent data points from well-isolated single units.

sample, which had an approximately Gaussian distribution of sharpness of tuning. The experimentally induced bias toward sharper tuning among MGBv-A1-like units presumably would have introduced a bias toward stronger, higher-acuity SSS among MGBv-A1-like units compared with A1. The greater across-unit variation in SSS among A1 units compared with MGBv-A1 units is evident in distributions shown in Figures 6 and 7.

The trends in distributions of SSS magnitude among neural populations and base rates were largely constant across ranges of frequency tuning of units, represented by their CFs. We tested for a correlation between the accuracy of SSS and unit CF by performing a Spearman rank-correlation analysis with 10,000 bootstrapped replications. SSS accuracy was taken as the greatest  $d'$  magnitude across all A-B separations of 20° to the left or right of A sources at C40°, 0°, and 140° (i.e., maximum  $d'$  across 6 A-B locations, all within 20° within A and B). For each replication, we randomly drew with replacement an equal number of units per one-octave CF bin from each unit population and across all bps conditions. CIs were calculated from each distribution of correlation coefficients (empirical two-tailed). We found a weak but significant positive relationship between  $d'$  and CF among BIN units at 5 bps (correlation coefficient CI = [0.08, 0.92],  $p < 0.05$ ; Spearman-rank correlation) and MGBv-ICC-like units across 10 (CI = [0.18, 0.95]), 20 (CI = [0.04, 0.91]), and 40 (CI = [0.13, 0.89]) bps ( $p < 0.05$ ; Spearman-rank correlation), which we regard as of little practical importance. No significant CF dependence of  $d'$  was seen among A1, MGBv-A1-like, and ICC unit

populations across all tested bps conditions (CIs = [−0.11 to −0.87, 0.34 to 0.95],  $p > 0.05$ ; Spearman-rank correlation).

### Contribution of forward suppression to SSS

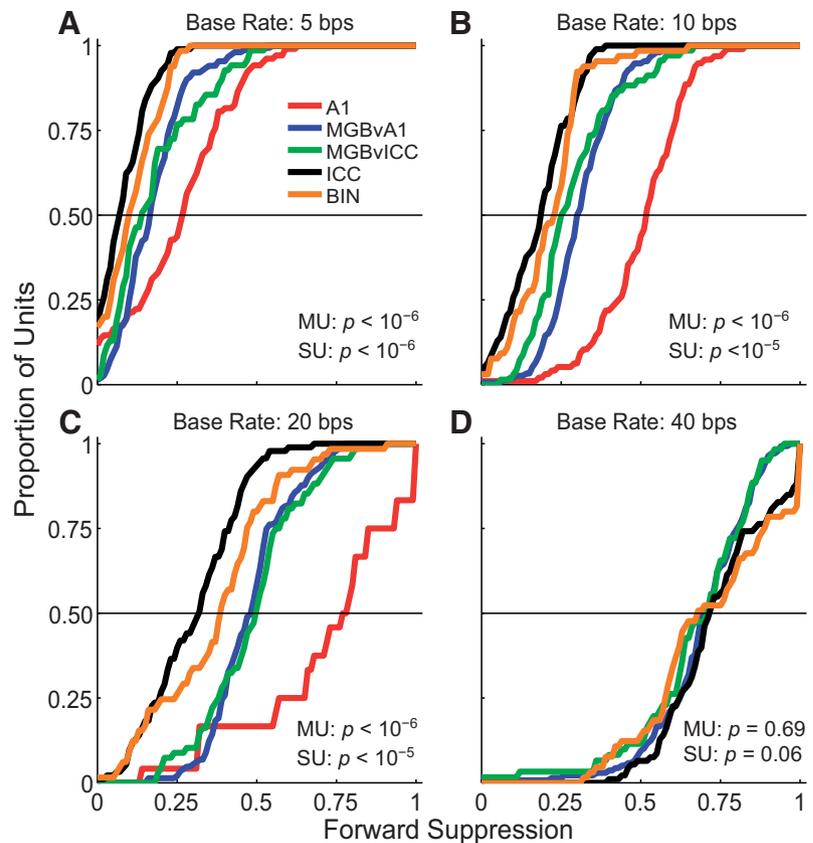
Many neurons, particularly in A1, showed a substantial decrease in spikes per sound burst under conditions of collocated A and B sources at C40° compared with a single source at C40°; we refer to this as “forward suppression” and a neutral term that could encompass a number of mechanisms, including refractoriness, forward inhibition, and/or synaptic depression. This can be seen in the example shown in Figure 3E. For all units, we quantified the amount of forward suppression by the fractional reduction in spikes per sound burst between single source and collocated competing sources at C40°; values approaching 0 and 1 indicate weak and strong forward suppression, respectively. The cumulative distributions of forward suppression for all unit populations at each bps condition are shown in Figure 8. At base rates of 5, 10, and 20 bps, we found a significant difference in forward suppression across all unit populations (MU:  $\chi^2 = 66.1$ –195.2,  $p < 10^{-6}$ ; SU:  $\chi^2 = 24.9$ –32.3,  $p < 10^{-5}$ ; Kruskal–Wallis). *Post hoc* comparisons indicated that forward suppression was greatest among A1 units at all base rates at which A1 units were tested, 5–20 bps ( $p < 0.0001$ , Bonferroni-corrected). The stronger forward suppression among A1 units was also evident in their reduced capability to synchronize to trains of sound bursts from single sources. Figure 9A plots the median values of vector strength for phase locking to trains of bursts at 2.5–40 bps, and Figure 9B plots

the percentage of neurons showing statistically significant phase locking to those rates. By both measures, it is clear that the ability of neurons to synchronize to sequences of noise bursts is dramatically decreased between subcortical regions and A1.

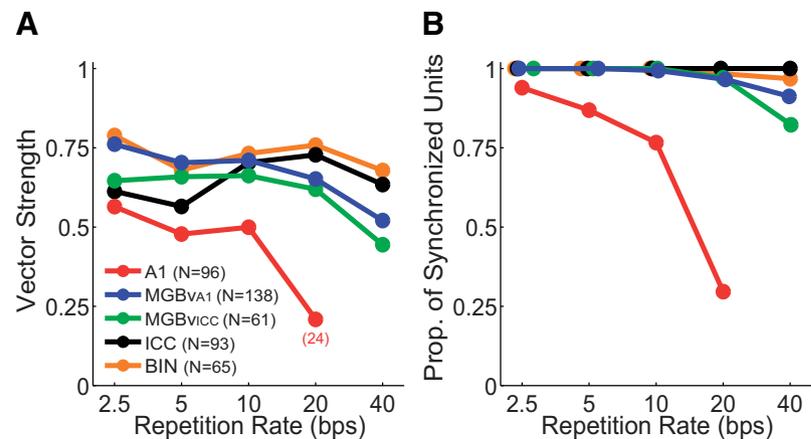
The addition of a competing source tended to sharpen the spatial tuning of units. This effect tended to increase with increasing stimulus base rate and was stronger among A1 units than among subcortical units. We quantified the sharpening of spatial tuning by calculating the ERRF width (defined in Materials and Methods) for each unit under single- and competing-source conditions (Fig. 10). The reduction in ERRF width (in degrees) between conditions varied significantly across unit populations at all base rates (MU and SU:  $\chi^2 = 30.4\text{--}299.7$ ,  $p < 0.001$ ; Kruskal–Wallis). *Post hoc* analysis showed that addition of a competing sound produced substantial sharpening among A1 units at all tested base rates ( $p < 0.0001$ , Bonferroni-corrected), whereas considerable sharpening was only evident among ICC, BIN, and MGBv units at faster base rates. This accords with the observation that only A1 units showed substantial forward suppression at base rates as low as 20 bps and suggests that A1 units, but not subcortical units, show a sharpening of spatial tuning that includes a major contribution from forward suppression.

### Forward suppression in A1 is not due to synaptic inhibition

Our measures of spatial stream segregation at multiple levels of the auditory pathway demonstrate a dramatic increase in forward suppression between subcortical regions and A1. We hypothesized that forward suppression represents either synaptic inhibition within the cortex or some other biophysical property of A1 neurons that limits the following rates in A1. We explored the putative inhibitory mechanism by recording extracellular neural responses from A1 neurons while applying GABA antagonists to the cortical surface (see Experimental procedures). Three GABA antagonists were used as follows: (1) Gabazine, an antagonist of postsynaptic GABA<sub>A</sub> inhibition; (2) CGP 36216, an antagonist of presynaptic GABA<sub>B</sub> inhibition; and (3) 2-Hydroxysaclofen, an antagonist of postsynaptic GABA<sub>B</sub> inhibition. We measured responses to pulse train stimuli presented at various repetition rates before and after drug application. The repetition rate cutoff was taken as the maximum repetition rate (Hz) at which responses were  $\geq 50\%$  of the maximum response across all tested repetition rates. If forward suppression in A1 was due to synaptic inhibition, we would expect application of

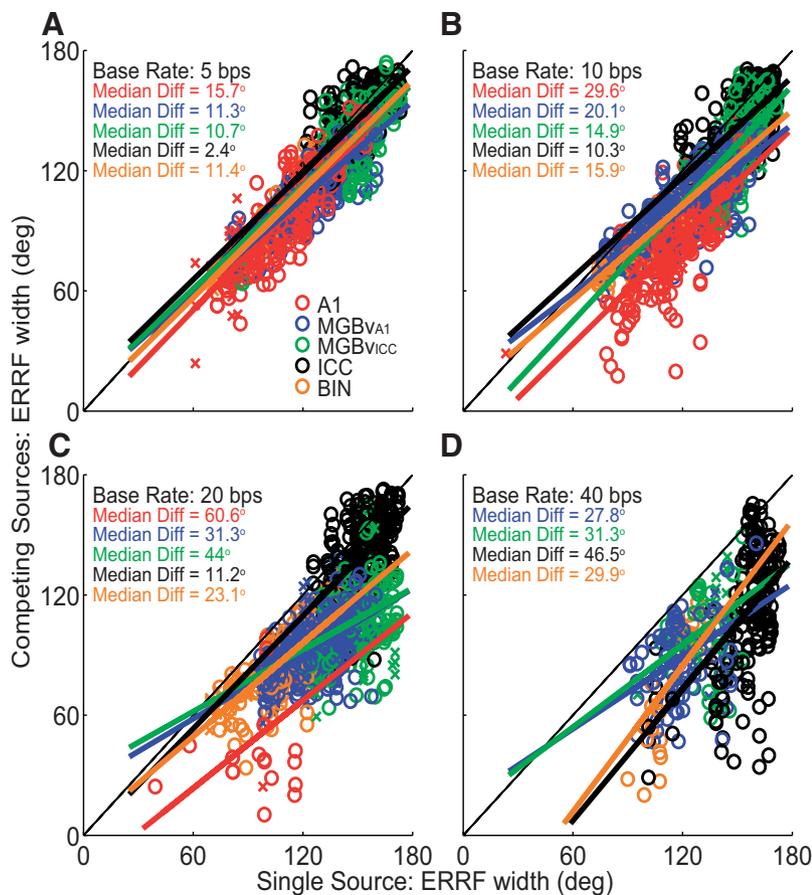


**Figure 8.** Cumulative distribution of forward suppression across each unit population (color coded) for each tested base rate (**A**, 5 bps; **B**, 10 bps; **C**, 20 bps; **D**, 40 bps). Statistics from a nonparametric ANOVA (Kruskal–Wallis) test across the multiunit and single-unit populations are shown in each subplot.



**Figure 9.** Median vector strength (**A**) and the proportion of synchronized units (**B**) for each unit population (color coded) at varying repetition rates (bps; horizontal axis). *x*-coordinates are slightly offset for visual purposes. Total number (N) of each unit population is shown. Only 24 A1 units were tested at the 20 bps rate.

GABA antagonists to lead to an increase in the stimulus repetition rate to which A1 neurons synchronize. In addition, the targeted receptor specificity of the agents would potentially indicate the source of intracortical synaptic inhibition. Surprisingly, we found that none of the GABA antagonists produced the hypothesized relief from forward suppression (Fig. 11A–C). Specifically, no significant change in repetition rate cutoffs was seen between pre-application and post-application of any of the three agents



**Figure 10.** Breadth of spatial sensitivity for the single-source (horizontal axis) versus the competing-source (vertical axis) condition across all unit populations for each tested base rate (**A**, 5 bps; **B**, 10 bps; **C**, 20 bps; **D**, 40 bps). Data points below the diagonal line indicate sharper spatial tuning in the presence of a competing sound. Circles represent multiple units.  $\times$ , Single units. Diagonal black line indicates the unity line. Thick colored lines indicate linear regression fits for corresponding ERRF widths for each unit population.

(Gabazine,  $p = 0.30$ ; CGP 36216,  $p = 0.20$ ; 2-Hydroxysaclofen,  $p = 0.36$ ; signed-rank tests). A similar lack of effect on repetition rate cutoff was seen at all cortical depths. That the drugs reached cortical neurons in effective concentrations was demonstrated by an overall increase in spikes per burst (Fig. 11D; Gabazine,  $t_{(61)} = 6.47$ ,  $p < 10^{-6}$ ; CGP 36216,  $t_{(46)} = 7.86$ ,  $p < 10^{-6}$ ; 2-Hydroxysaclofen,  $t_{(58)} = 5.46$ ,  $p < 10^{-6}$ ; paired  $t$  tests); again, this was seen at all cortical depths. These results are inconsistent with an explanation for forward suppression based on intracortical inhibition.

## Discussion

We evaluated SSS at three levels of the ascending auditory system: midbrain (ICC and BIN), thalamus (two subpopulations of MGBv neurons), and cortical area A1. The results demonstrate that the degree to which neurons preferentially synchronize to sounds from one or the other of two sources is progressively enhanced along the ascending pathway, with robust SSS observed in essentially all A1 neurons that we tested. The enhancement of SSS reflects the sharpened spatial sensitivity and strengthened forward suppression at higher levels of the auditory pathway. Moreover, we found that forward suppression within the auditory cortex was not due to intracortical synaptic inhibition.

### Stream segregation along the ascending auditory system

Physiological studies of stream segregation based on differences in tone frequencies have demonstrated neural correlates within

the mammalian auditory cortex (Fishman et al., 2001, 2004, 2012; Michey et al., 2005; Elhilali et al., 2009) and avian forebrain (Bee and Klump, 2004; Bee et al., 2010). Stream segregation based on differences in spatial location has been identified in the responses of cortical neurons (Middlebrooks and Bremen, 2013; present study), and physiological correlates of stream segregation based on interaural time differences are observed in the auditory cortex of human listeners in studies using magnetoencephalography (Carl and Gutschalk, 2013) and fMRI (Schadwinkel and Gutschalk, 2010). Other studies have reported that stream segregation is present at subcortical levels, with tone-based stream segregation seen among single-unit responses from the cochlear nucleus (Pressnitzer et al., 2008), and interaural time difference-based behavioral streaming linked with fMRI BOLD activity in the IC (Schadwinkel and Gutschalk, 2011). These reports, together with the present results, offer the view that some forms of stream segregation can be present throughout all levels of the ascending auditory pathway. Specifically with regard to spatial factors, however, we find that SSS begins with gradual sharpening of spatial sensitivity at successive levels of the brainstem and thalamus, and that SSS is enhanced by forward suppression between thalamic and cortical levels.

Our results are consistent with the failure to demonstrate location-based stream segregation at the level of the IC in guinea pigs (Shackleton et al., 2012). Although we encountered SSS among BIN and a subpopulation of MGBv neurons at faster base rates (Fig. 4), those presentation rates are considerably faster than the timescale reported in psychophysical studies of stream segregation. Schadwinkel and Gutschalk's (2011) findings of associated fMRI BOLD activity in the IC with interaural time difference-based streaming might be attributed to active engagement or attentional modulation. Kondo and Kashino (2009) report that feedforward and feedback processes along the thalamocortical loop are involved in the formation of auditory streaming percepts. Thus, further work should attempt to distinguish the roles of corticothalamic and corticotectal modulation that aid in auditory streaming.

### Segregating streams through spatial hearing and forward suppression

In our rat animal model, spatial sensitivity develops along the ascending tectal and lemniscal pathway, from level-dependent spatial sensitivity that broadens markedly with increasing sound levels within the ICC to level-tolerant contralateral hemifield spatial sensitivity within the BIN, a subpopulation of neurons in MGBv, and area A1 (Yao et al., 2015). Not surprisingly, SSS was most prominent among A1, MGBv-A1-like, and BIN units. BIN and MGBv-A1-like units displayed SSS by virtue of their dominant contralateral hemifield tuning, whereas SSS among A1 units was due to contralateral hemifield tuning enhanced by forward

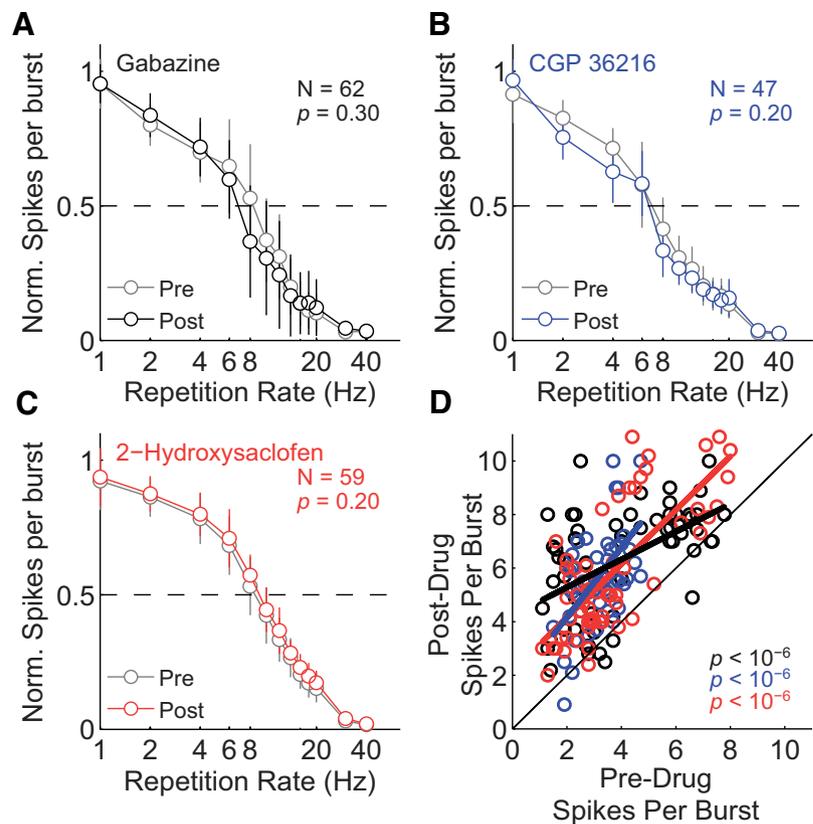
suppression at that level. These results give further evidence for two parallel pathways for auditory space processing: tectal, projecting to the superior colliculus; and lemniscal, projecting to the forebrain (Knudsen et al., 1993; Knudsen and Knudsen, 1996; Yao et al., 2015). Whether or not the representation of segregated streams along the tectal pathway plays a role in auditory scene analysis remains to be tested.

Despite the species differences in single-source spatial sensitivity, we encountered SSS results from A1 neurons in the anesthetized rat very similar to those observed among cortical neurons in anesthetized cats (Middlebrooks and Bremen, 2013). Specifically, we found that segregation of competing sources within cortical neurons was weakest when both sources were located within the contralateral hemifield and strongest when one of the two sources was located in the ipsilateral hemifield, as shown in Figures 3 and 7. This indicates that a cortical neuron's spatial sensitivity, which is derived from binocular computations within the brainstem and likely inherited from MGBv-A1-like units (Kyweriga et al., 2014), favors one of the two competing sound sources. That spatial bias is amplified by additional forward suppression. In particular, A1 neurons show strong suppression of responses in the condition when competing sounds are colocated and a strong release from suppression when one source is moved away from the other, yielding neural responses that are captured by one source. Our interpretation for such findings is that SSS begins in a subpopulation of neurons within the MGBv and is further enhanced along the thalamocortical synapse, becoming dominant at the cortical level.

#### Potential mechanisms of forward suppression

Consistent with previous reports, our measures of SSS at multiple levels of the auditory pathway demonstrate a dramatic increase in forward suppression and corresponding decrease in upper rate cutoffs for synchrony to repeated stimuli between the MGBv and A1. Our results accord with observations of forward suppression in tone-based streaming studies (Fishman et al., 2001, 2004; Bee and Klump, 2004) and with measures of spatially dependent forward suppression with leading and lagging sounds (Reale and Brugge, 2000; Mickey and Middlebrooks, 2005; Zhou and Wang, 2014). Also, the timescale of forward suppression that we measured in the context of SSS agrees with that of suppression observed in forward masking studies where the response to a probe stimulus is largely suppressed following a preceding masker stimulus (Calford and Semple, 1995; Brosch and Schreiner, 1997; Scholes et al., 2011). These findings suggest that similar cortical mechanisms could be involved in forward masking, forward suppression of temporal sequences, and segregating sequential sounds into discrete streams.

We used a pharmacological procedure to test the hypothesis that forward suppression in the auditory cortex is due to synaptic



**Figure 11.** Effects of GABA<sub>A</sub> or GABA<sub>B</sub> antagonists on A1 responses. Normalized grand mean spikes per burst ( $\pm 1$  SD) to pulse train stimuli presented at varying repetition rates (horizontal axis) are plotted before and after application of Gabazine (**A**), CGP 36216 (**B**), and 2-Hydroxysaclofen (**C**). **D**, Mean spikes per burst of each unit for pre-drug (horizontal axis) and post-drug (vertical axis) application. *p* values indicate results from comparing the mean repetition rate cutoffs (signed-rank test) (**A–C**) or mean spikes per burst (**D**) between pre-drug and post-drug application samples (paired *t* test).

(GABAergic) inhibition. Interestingly, we found that cortical neurons did not display the hypothesized relief from forward suppression after drug application, suggesting that forward suppression is not due to synaptic inhibition. Results from Wehr and Zador (2005) indicate that synaptic inhibition plays a small role in forward suppression. In that study, they conducted whole-cell recordings on neurons in the rat auditory cortex and found that inhibitory postsynaptic potentials elicited by forward masking stimuli were brief, lasting no more than 100 ms. Thus, the forward suppression that we observed on a  $>100$  ms timescale could not result from the brief synaptic inhibition of cortical neurons. Our negative results against intracortical synaptic inhibition, in addition with the findings from Wehr and Zador's (2005) study, refute the synaptic inhibition hypothesis for forward suppression.

Forward suppression seen within A1 could reflect various mechanisms. One hypothesis involves the biophysical property of postdischarge adaptation (i.e., refractoriness). Middlebrooks and Bremen (2013), however, demonstrated that the probability of action potential firing in response to a sound was independent of firing elicited by a preceding sound. This was particularly the case under the 5 bps condition, with some indication of intracortical adaptation in the 10 bps condition. The 10 bps value accords well with the time constant of forward masking in A1, evident in Figure 9 and in previous reports (e.g., Creutzfeldt et al., 1980; Schreiner and Urbas, 1988). This argues against cortical postdischarge adaptation as a mechanism of forward suppression. Other potential sources of cortical forward suppression could be inher-

itance from thalamic inputs or synaptic depression at the thalamocortical synapse. It is unlikely that cortical neurons directly inherit their forward suppression from thalamic neurons because MGBv neurons can follow periodic stimuli at much higher repetition rates compared with their cortical inputs (Creutzfeldt et al., 1980). It is more likely that cortical forward suppression reflects synaptic depression of thalamocortical synapses. Findings from intracellular recordings suggest that the low-pass temporal filtering between the thalamic and cortical level is the result of an activity-dependent decrease in synaptic transmission (Varela et al., 1997; Chance et al., 1998; Fortune and Rose, 2000). Furthermore, results from computational modeling studies have demonstrated that cortical repetition rate suppression can be modeled by presynaptic depression and a small amount of facilitation (Eggermont, 1999, 2002). Recently, a study in mice by Bayazitov et al. (2013) demonstrated that synaptic depression along the thalamocortical synapse can explain the forward suppression seen in the auditory cortex. Specifically, they found that paired-pulse synaptic depression at thalamocortical projection sites is due to a switch between firing modes of thalamic neurons, which is dependent on Ca<sub>v</sub>3.1 T-type calcium channels. Pharmacologically inhibiting or RNA-mediated knockdown of those calcium channels significantly diminished synaptic depression at thalamocortical projections and forward suppression in the auditory cortex.

Based on the available reports, we hypothesize that cortical SSS is due to synaptic depression at the thalamocortical synapse. We hope to directly explore the synaptic depression hypothesis in future experiments.

## References

- Bartlett EL, Wang X (2007) Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *J Neurophysiol* 97:1005–1017. [CrossRef Medline](#)
- Bayazitov IT, Westmoreland JJ, Zakharenko SS (2013) Forward suppression in the auditory cortex is caused by the Ca(v)3.1 calcium channel-mediated switch from bursting to tonic firing at thalamocortical projections. *J Neurosci* 33:18940–18950. [CrossRef Medline](#)
- Bee MA, Klump GM (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. *J Neurophysiol* 92:1088–1104. [CrossRef Medline](#)
- Bee MA, Micheyl C, Oxenham AJ, Klump GM (2010) Neural adaptation to tone sequences in the songbird forebrain: patterns determinants, and relation to the build-up of auditory streaming. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 196:543–557. [CrossRef Medline](#)
- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Massachusetts Institute of Technology.
- Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol* 77:923–943. [Medline](#)
- Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. *J Neurophysiol* 73:1876–1891. [Medline](#)
- Carl D, Gutschalk A (2013) Role of pattern, regularity, and silent intervals in auditory stream segregation based on interaural time differences. *Exp Brain Res* 224:557–570. [CrossRef Medline](#)
- Chance FS, Nelson SB, Abbott LF (1998) Synaptic depression and the temporal response characteristics of V1 cells. *J Neurosci* 18:4785–4799. [Medline](#)
- Creutzfeldt O, Hellweg FC, Schreiner C (1980) Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 39:87–104. [Medline](#)
- Doron NN, Ledoux JE, Semple MN (2002) Redefining the tonotopic core of rat auditory cortex: physiological evidence for a posterior field. *J Comp Neurol* 453:345–360. [CrossRef Medline](#)
- Eggermont JJ (1999) The magnitude and phase of temporal modulation transfer functions in cat auditory cortex. *J Neurosci* 19:2780–2788. [Medline](#)
- Eggermont JJ (2002) Temporal modulation transfer functions in cat primary auditory cortex: separating stimulus effects from neural mechanisms. *J Neurophysiol* 87:305–321. [Medline](#)
- Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* 61:317–329. [CrossRef Medline](#)
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear Res* 151:167–187. [CrossRef Medline](#)
- Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *J Acoust Soc Am* 116:1656–1670. [CrossRef Medline](#)
- Fishman YI, Micheyl C, Steinschneider M (2012) Neural mechanisms of rhythmic masking release in monkey primary auditory cortex: implications for models of auditory scene analysis. *J Neurophysiol* 107:2366–2382. [CrossRef Medline](#)
- Fortune ES, Rose GJ (2000) Short-term synaptic plasticity contributes to the temporal filtering of electrosensory information. *J Neurosci* 20:7122–7130. [Medline](#)
- Gaese BH, Ostwald J (1995) Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *Eur J Neurosci* 7:438–450. [CrossRef Medline](#)
- Games KD, Winer JA (1988) Layer V in rat auditory cortex: projections to the inferior colliculus and contralateral cortex. *Hear Res* 34:1–25. [CrossRef Medline](#)
- Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32:613–636. [Medline](#)
- Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.
- Heaulme M, Chabon JP, Leyris R, Molimard JC, Wermuth CG, Biziere K (1986) Biochemical characterization of the interaction of three pyridazinyl-GABA derivatives with the GABA<sub>A</sub> receptor site. *Brain Res* 384:224–231. [CrossRef Medline](#)
- Heffner HE, Heffner RS (1985) Sound localization in wild Norway rats (*Rattus norvegicus*). *Hear Res* 19:151–155. [CrossRef Medline](#)
- Heffner HE, Heffner RS, Contos C, Ott T (1994) Audiogram of the hooded Norway rat. *Hear Res* 73:244–247. [CrossRef Medline](#)
- Higgins NC, Storace DA, Escabi MA, Read HL (2010) Specialization of binaural responses in ventral auditory cortices. *J Neurosci* 30:14522–14532. [CrossRef Medline](#)
- Ito M, van Adelle B, Kelly JB (1996) Sound localization after transection of the commissure of Probst in the albino rat. *J Neurophysiol* 76:3493–3502. [Medline](#)
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577. [CrossRef Medline](#)
- Kavanagh GL, Kelly JB (1986) Midline and lateral field sound localization in the albino rat (*Rattus norvegicus*). *Behav Neurosci* 100:200–205. [CrossRef Medline](#)
- Kelly JB, Masterton B (1977) Auditory sensitivity of the albino rat. *J Comp Physiol Psychol* 91:930–936. [CrossRef Medline](#)
- Kerr DI, Ong J, Johnston GA, Abbenante J, Prager RH (1988) 2-Hydroxy-saclofen: an improved antagonist at central and peripheral GABA<sub>B</sub> receptors. *Neurosci Lett* 92:92–96. [CrossRef Medline](#)
- Kilgard MP, Merzenich MM (1999) Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hear Res* 134:16–28. [CrossRef Medline](#)
- Knudsen EI, Knudsen PF (1996) Disruption of auditory spatial working memory by inactivation of the forebrain archistriatum in barn owls. *Nature* 383:428–431. [CrossRef Medline](#)
- Knudsen EI, Knudsen PF, Masino T (1993) Parallel pathways mediating both sound localization and gaze control in the forebrain and midbrain of the barn owl. *J Neurosci* 13:2837–2852. [Medline](#)
- Kondo HM, Kashino M (2009) Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. *J Neurosci* 29:12695–12701. [CrossRef Medline](#)
- Kotak VC, Takesian AE, Sanes DH (2008) Hearing loss prevents the maturation of GABAergic transmission in the auditory cortex. *Cereb Cortex* 18:2098–2108. [CrossRef Medline](#)
- Kyweriga M, Stewart W, Cahill C, Wehr M (2014) Synaptic mechanisms underlying interaural level difference selectivity in rat auditory cortex. *J Neurophysiol* 112:2561–2571. [CrossRef Medline](#)

- Lee CC, Middlebrooks JC (2011) Auditory cortex spatial sensitivity sharpens during task performance. *Nat Neurosci* 14:108–114. [CrossRef Medline](#)
- Macmillan NA, Creelman CD (2005) *Detection theory: a user's guide*, Ed 2. Mahwah, NJ: Lawrence Erlbaum.
- Mardia KV (1972) *Statistics of directional data*. London: Academic.
- Metherate R, Ashe JH (1994) Facilitation of an NMDA receptor-mediated EPSP by paired-pulse stimulation in rat neocortex via depression of GABAergic IPSPs. *J Physiol* 481:331–348. [CrossRef Medline](#)
- Micheyl C, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48:139–148. [CrossRef Medline](#)
- Mickey BJ, Middlebrooks JC (2005) Sensitivity of auditory cortical neurons to the locations of leading and lagging sounds. *J Neurophysiol* 94:979–989. [CrossRef Medline](#)
- Middlebrooks JC, Bremen P (2013) Spatial stream segregation by auditory cortical neurons. *J Neurosci* 33:10986–11001. [CrossRef Medline](#)
- Middlebrooks JC, Onsan ZA (2012) Stream segregation with high spatial acuity. *J Acoust Soc Am* 132:3896–3911. [CrossRef Medline](#)
- Ong J, Bexis S, Marino V, Parker DA, Kerr DI, Froestl W (2001) CGP 36216 is a selective antagonist at GABA(B) presynaptic receptors in rat brain. *Eur J Pharmacol* 415:191–195. [CrossRef Medline](#)
- Polley DB, Read HL, Storaice DA, Merzenich MM (2007) Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *J Neurophysiol* 97:3621–3638. [CrossRef Medline](#)
- Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual organization of sound begins in the auditory periphery. *Curr Biol* 18:1124–1128. [CrossRef Medline](#)
- Reale RA, Brugge JF (2000) Directional sensitivity of neurons in the primary auditory (AI) cortex of the cat to successive sounds ordered in time and space. *J Neurophysiol* 84:435–450. [Medline](#)
- Rees A, Møller AR (1987) Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear Res* 27:129–143. [CrossRef Medline](#)
- Rutkowski RG, Miasnikov AA, Weinberger NM (2003) Characterisation of multiple physiological fields within the anatomical core of rat auditory cortex. *Hear Res* 181:116–130. [CrossRef Medline](#)
- Schadwinkel S, Gutschalk A (2010) Activity associated with stream segregation in human auditory cortex is similar for spatial and pitch cues. *Cereb Cortex* 20:2863–2873. [CrossRef Medline](#)
- Schadwinkel S, Gutschalk A (2011) Transient bold activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. *J Neurophysiol* 105:1977–1983. [CrossRef Medline](#)
- Scholes C, Palmer AR, Sumner CJ (2011) Forward suppression in the auditory cortex is frequency-specific. *Eur J Neurosci* 33:1240–1251. [CrossRef Medline](#)
- Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat: II. Comparison between cortical fields. *Hear Res* 32:49–63. [CrossRef Medline](#)
- Shackleton T, Middlebrooks J, Palmer A (2012) Spatial position related adaptation in the inferior colliculus of anesthetized guinea pigs. *Abs Assoc Res Otolaryngol* 35.
- Varela JA, Sen K, Gibson J, Fost J, Abbott LF, Nelson SB (1997) A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J Neurosci* 17:7926–7940. [Medline](#)
- Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. *Neuroscience* 157:484–494. [CrossRef Medline](#)
- Wehr M, Zador AM (2005) Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron* 47:437–445. [CrossRef Medline](#)
- Yao JD, Bremen P, Middlebrooks JC (2013) Rat primary auditory cortex is tuned exclusively to the contralateral hemifield. *J Neurophysiol* 110:2140–2151. [CrossRef Medline](#)
- Yao JD, Bremen P, Middlebrooks JC (2015) Transformation of spatial sensitivity along the ascending auditory pathway. *J Neurophysiol* 113:3098–3111. [CrossRef Medline](#)
- Zhou B, Green DM, Middlebrooks JC (1992) Characterization of external ear impulse responses using Golay codes. *J Acoust Soc Am* 92:1169–1171. [CrossRef Medline](#)
- Zhou Y, Wang X (2014) Spatially extended forward suppression in primate auditory cortex. *Eur J Neurosci* 39:919–933. [CrossRef Medline](#)