SUPPLEMENTAL DATA

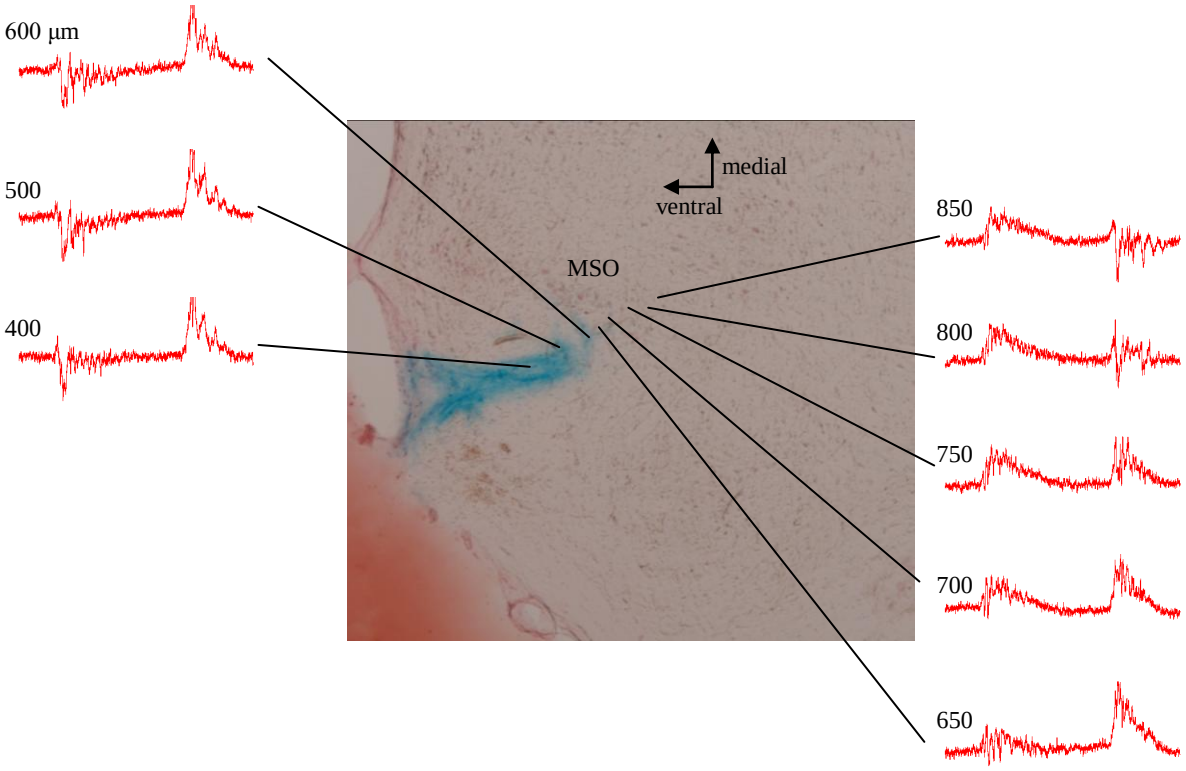

Figure S1 (related to Figure 1). Identification of the somatic layer in the gerbil MSO using a ventral approach.
Section of the ventral brainstem with the electrode track marked in blue by pressure ejection of Alcian Blue. Red traces show the local field potential response ('neurophonics') during alternating monaural click stimuli to the left and right ear (Biedenbach and Freeman, 1964; Clark and Dunlop, 1968; Galambos et al., 1959). Numbers indicate penetration depth from surface. The somatic layer ('MSO') lies between the point where the response to contralateral tones reverses (~650 µm) and the point where the ipsilateral click response reverses (~850 µm).

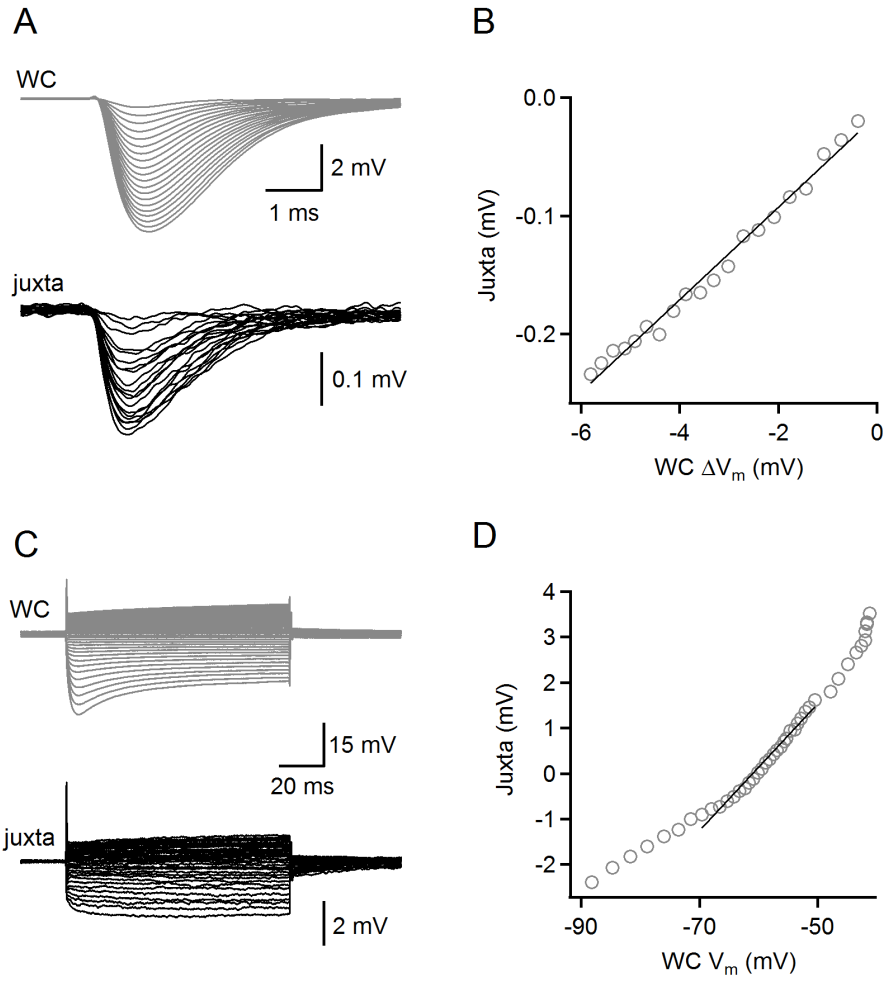

Figure S2 (related to Figure 2). Relation between juxtacellular and whole-cell potentials is linear for IPSPs and for current steps within the physiological membrane potential range.

(A) Sample whole cell and juxtacellular responses to IPSPs generated by conductance clamp simulation of an IPSG over a range of 20 conductance levels. Data are averages of 20 trials. (B) Linear correlation between peak IPSP amplitudes measured with juxtacellular and whole cell recordings from the cell shown in (A) (r = 0.994). (C) Sample whole cell and juxtacellular responses to current steps ranging from -5000 to 5000 pA in 200 pA increments. Data are averages of 6 trials. (D) Over the physiologically important voltage range (-70 to -50 mV), peak responses to currents steps were linearly correlated between juxtacellular and whole cell recordings from the cell in (C) (r = 0.989).

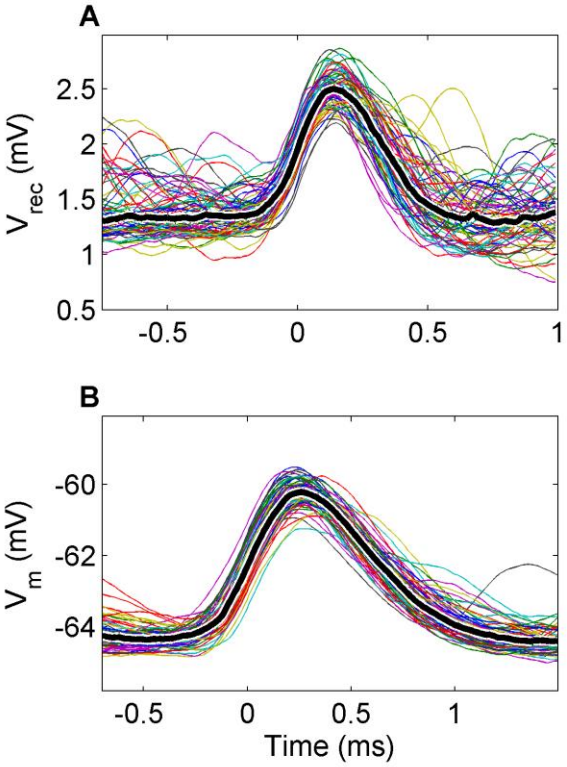

Figure S3 (related to Figure 2). Absence of downward-going phase preceding EPSPs. Subthreshold events were extracted from the MSO recordings based on their steep upward slopes. Each panel shows the 70 largest events (colored thin lines) and the median value (thick black line) obtained from a different MSO neuron. The events are aligned on their steepest rising phase (time zero). (A) Juxtacellular recording; CF = 300 Hz ; (B) Whole-cell recording; CF = 780 Hz.

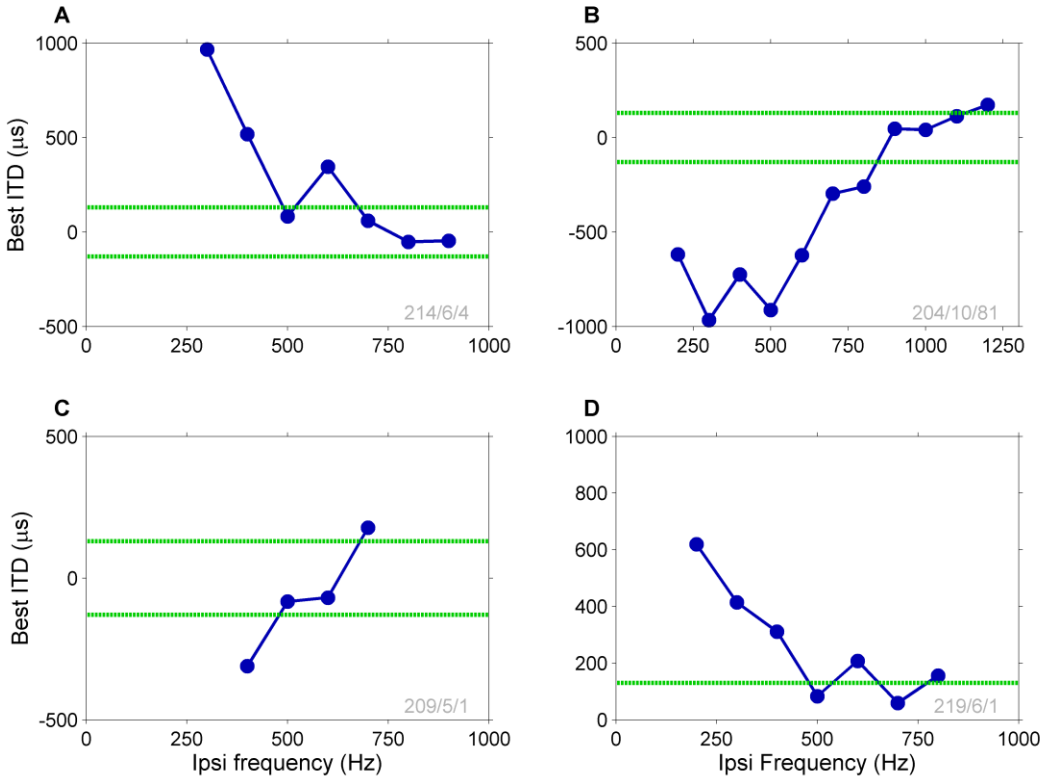

Figure S4 (related to Figure 3). Variation of ITD with stimulus frequency.
The best ITD of many MSO neurons varied considerably with stimulus frequency. Best ITD was evaluated from the peak of the AP cycle histogram derived from responses to 60-dB-SPL, 4-Hz binaural beat stimuli. The four panels show data from four different MSO neurons. Horizontal green lines mark the ±130 µs physiological range of ITDs determined by the head size of the gerbil. CFs: 680, 760, 560, 330 Hz.

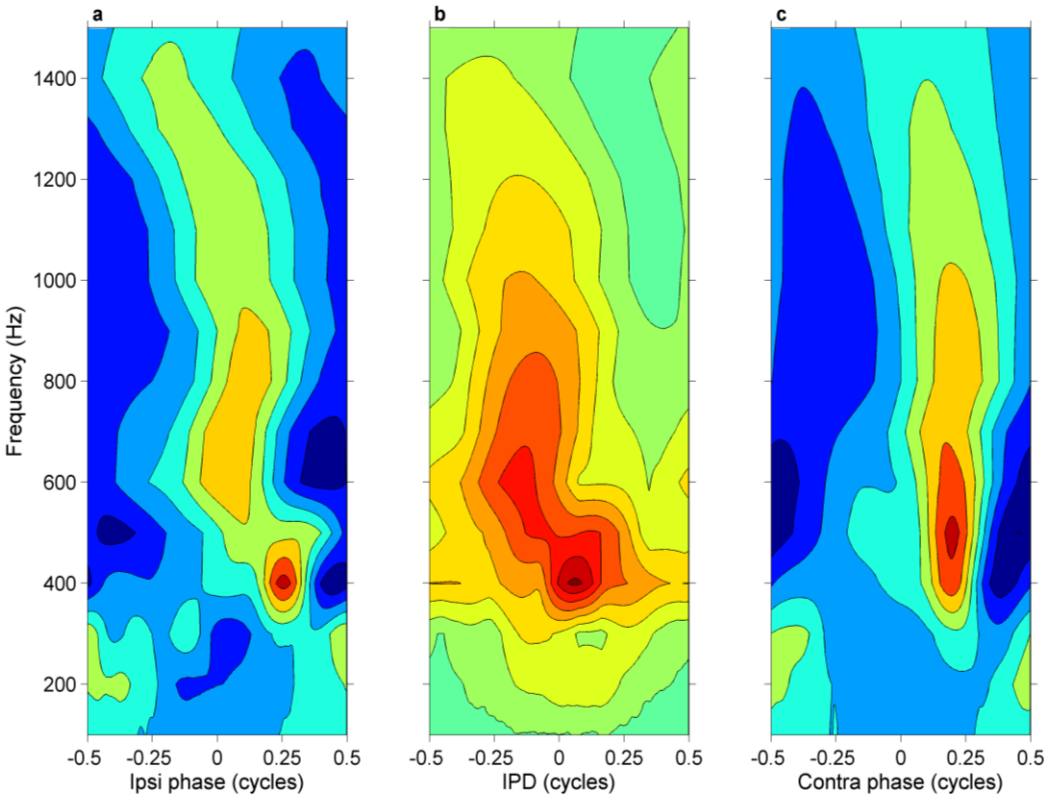

Figure S5 (related to Figure 4). Frequency-dependent monaural and binaural tuning. (A) and (C) show the monaural cycle-averaged waveforms (a subset of which is shown in Figure 4A) in the form of contour plots (0.2-mV spacing) as a function of both stimulus frequency and the monaural phase relative to the stimulus onset. Monaural phases were compensated for a 5.5-ms overall stimulus-response latency delay estimated from the onset latency of the recording. The subthreshold is phase-locked over the entire range of frequencies, reflected by systematic vertical "ridges." The steep transition of the monaural phase pattern between 300 Hz and 400 Hz produces a complex phase-frequency relation, which, however, is matched across the two ears. (B) shows the reduction of these monaural phase data to the IPD axis (cf. the green line in Figures 3A and 4B). The resulting "binaural receptive field" shows the tuning of the subthreshold input as a function of frequency and IPD. Contour spacing, 0.1 mV. The interaural matching of the monaural transition causes the binaural receptive field to be more regular than the monaural phase-frequency plots in A and C. The binaural tuning is still complex, and cannot be captured by a single, frequency-independent delay, because that would cause the ridge to be on a straight line through the origin (zero IPD and zero frequency) See also Day and Semple (2011). CF: 680 Hz.

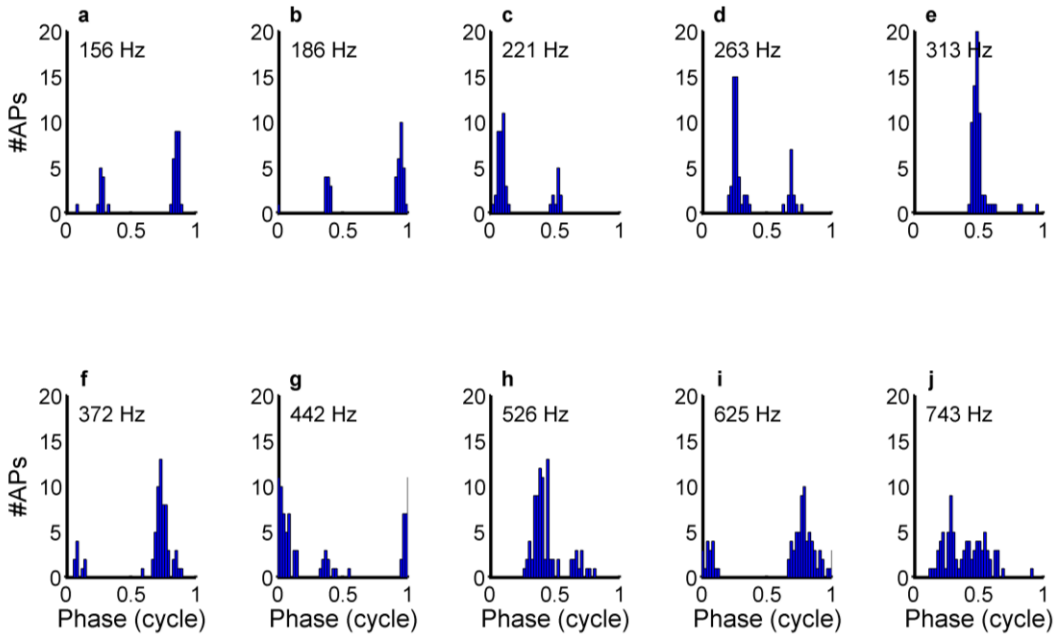

Figure S6 (related to Figure 4). Multiple peaks in cycle histograms of spherical bushy cells (SBCs) of the gerbil.
Unpublished data from unit RG0981/1u1 of Kuenzel et al. (2011). Cycle histograms of APs in response to 50-ms, 80-dB-SPL single tones, analyzed as described in Kuenzel et al. (2011). Such "peak splitting", was a common observation in the SBC responses to low-frequency (<500 Hz) tones at intensities of 50-80 dB SPL. Courtesy of Dr. Thomas Kuenzel, Aachen.

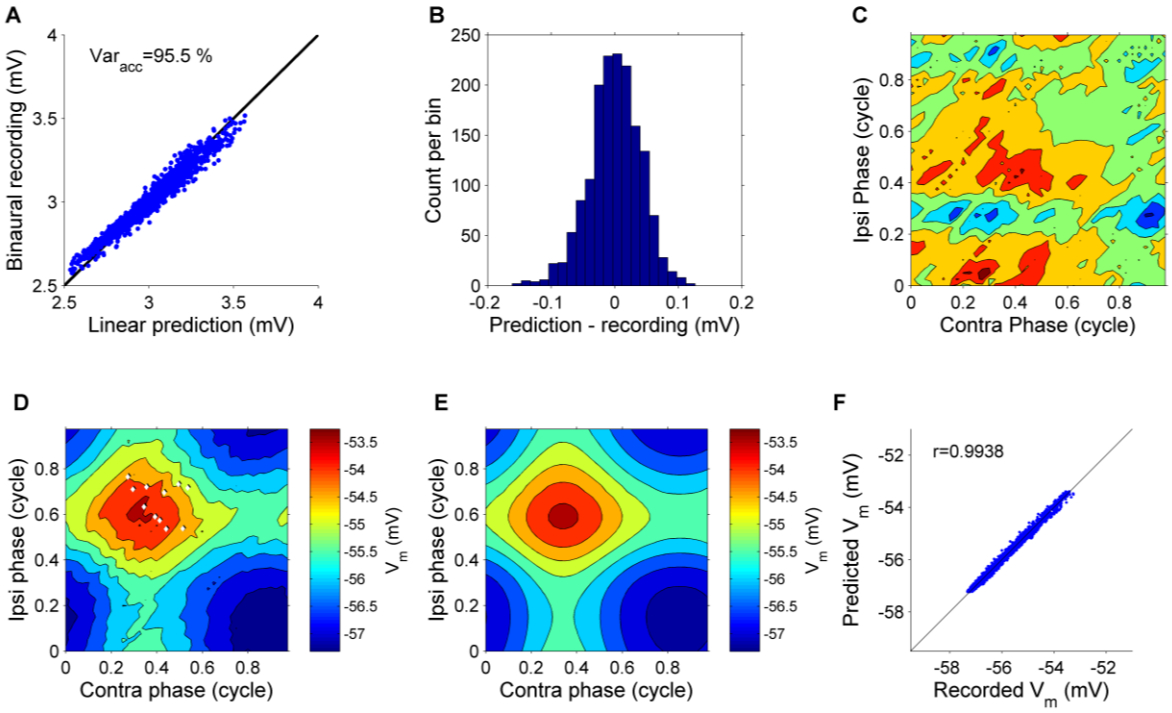

Figure S7 (related to Figure 5). Linear summation of monaural inputs.
(A) Scatter plot comparing linear prediction of the subthreshold input obtained from summing monaural recordings (Figure 5D) with the actual binaural response (Figure 5A). CF: 680 Hz. The black line indicates unity. The linear prediction accounted for 95.5% of the variance in the data. (B) Histogram of the deviations, i.e. data minus prediction. (C) Contour plot of the deviations. Blue patches correspond to data points below the predictions. Contour spacing 0.05 mV. The trend of the predictions to overestimate the highest values (A) the skewness of -0.32 of the distribution (B), and the concentration of negative deviations (blue patches in C) on the peaks of the subthreshold response (compare to Figure 4A) are all consistent with a contribution of a small bias toward smaller EPSPs in the binaural recording caused by the removal of those events that actually triggered APs. The removed segments added up to 0.076% of the total duration of the waveform analyzed. (D) 2-D representation of subthreshold input as in Fig. 2, but now for an in vivo whole cell recording. Stimulus was a 6-s, 700/704-Hz binaural beat at 40 dB SPL. *White dots* indicate action potentials. (E) The prediction from linear summation of the monaural contributions. (F) The linear prediction plotted against the recorded subthreshold response, with the correlation coefficient *r* indicated in the graph.

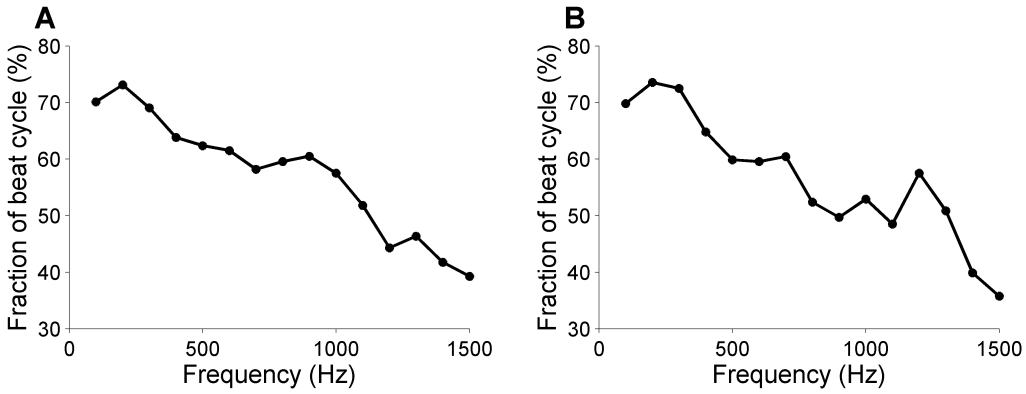

Figure S8 (related to Figure 6). Effect of frequency on the periodic reduction of subthreshold input.

(A) In response to 70-dB-SPL binaural beats having increasing monaural frequencies (*abscissa*), the dips in the variance of the subthreshold input became less pronounced, as reflected by the increasingly smaller fraction of the beat cycle during which the variance lies below the spontaneous value (*ordinate*). (B) A similar reduction of the fraction of below-spontaneous values of the variance was observed in the response to 70-dB-SPL monaural stimulation. The frequency dependence is in agreement with the assumption that the dips in the subthreshold potential are caused by the periodic absence of spikes from the monaural inputs; silent intervals will become less prominent for high stimulus frequencies, for which the phase-locking of the monaural inputs is known to degrade. CF: 680 Hz.

Movie S1 (related to Figure 5). Linear summation of monaural contributions predicts binaural subthreshold input.

The monaural contributions to the 400-Hz binaural beat stimulus were extracted by averaging over the respective stimulus periods of the two ears (cf. Figure 3C). The lower trace of the movie shows the ongoing ipsilateral (blue) and contralateral (red) inputs resulting from concatenating these single-cycle segments. The 4-Hz interaural frequency difference causes a running phase disparity between the inputs from the two ears. The upper trace shows the prediction (green) of the binaural subthreshold input obtained by the linear summation of the monaural waveforms. Its waxing and waning reflects the "beating" of the monaural inputs due to their slightly different periodicities. Along with the prediction the actual binaural subthreshold input, averaged over the 250-ms beat cycle, is shown (grey). The good match between the linear prediction and the measured response reveals that the response to the binaural beat can be well described as a beating between monaural inputs. The movie, which is best played in loop mode, provides a dynamic alternative to the static, two-dimensional mode of display of Figure 5A.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Data acquisition
Data were sampled at an interval of 10 µs with a 16-bit A/D converter (Digidata 1322A). Potentials were filtered at 10 kHz low-pass (four-pole Bessel filter). In 2 recordings data were additionally high-pass filtered at 1 Hz (one-pole Bessel filter). To enable online analysis during recordings, the recorded potentials were filtered (0.1 Hz to 30 kHz) and amplified (gain 10x) with a PARC 113 pre-amp. Spike and very large EPSPs were sorted based on amplitude of either a positive or negative peak with a discrimination system controller (BAK electronics). Online analysis was done in parallel with the MultiClamp recordings and revealed auditory tuning of cells to which the experiments could be specified.

Whole cell in vivo recordings
For whole-cell recordings, pipettes were advanced with positive pressure up to 10 µm until resistance approximately doubled, after which pressure was released, and suction was applied to obtain the cell-attached configuration. The whole-cell configuration was established by suction pulses. Whole-cell recordings were compensated for an estimated junction potential of -15 mV. Based on the fact that we only recorded from the somatic layer, as judged from the field potentials, most likely almost all juxtacellular and whole-cell recordings were somatic, although we cannot exclude that some recordings were made from proximal dendrites.

Analysis of in vivo recordings
The recordings were divided into 6500-ms or 9500-ms responses to single binaural-beat (or monaural tone) stimuli, each including a 500-ms spontaneous baseline immediately preceding stimulus onset. In order for a recorded response to be accepted for analysis, the RMS value of the 500-ms spontaneous portion had to exceed 0.1 mV and vary by no more than 25% across the individual responses by the same neuron. Recorded waveforms were filtered by a notch filter (50-Hz centre frequency; Q = 40, MATLAB function iirnotch) to remove any interference from AC power lines. Low-frequency artifacts caused by breathing and heart beat were attenuated by subtracting a low-pass filtered version of the waveform, which was obtained by downsampling to a sample interval of 250 ms, followed by upsampling to the original sample rate. This method of high-pass filtering ensures that frequency components at 4 Hz (the binaural beat frequency) and above are unaffected. Any effects on the cycle averages used in the subsequent analysis are thus excluded, while still reducing the spurious low-frequency variance. To avoid transient effects, the first 500 ms following stimulus onset (i.e., the first two beat cycles) were excluded from analysis.

The eAPs were identified based on the steep (~10 V/s) downward slopes following their peak. The time derivative of the recorded waveform was computed by convolution with a 50-µs-wide Hann window followed by subtracting adjacent samples. The local minima (within a moving 750-µs-long time window) were determined, and the resulting histograms (Figure 1D) were visually inspected. Only those recording showing a clear bimodal distribution of local slope minima were accepted for further analysis (19 out of 25 cells). Spike arrival times were obtained by determining the temporal position of the peaks immediately preceding the steep downward slope. The timing of the EPSPs was obtained by either determining the time of the peak immediately preceding the eAP peak, or, when no such peak was present in the 500 µs preceding the eAP, by finding the local minimum of the second time derivative ("inflection point") in that interval (Lorteije et al., 2009).

Cycle-averages (Figures 3C, D; 4A; 5A, C, D, F) were obtained by dividing the waveform into 1/*f*-long segments (*f* is the stimulus frequency tested, i.e., the ipsilateral or contralateral stimulus frequency or their 4-Hz difference) and computing the average across the corresponding time points of the segments. Any samples lying within 750 μs from an eAPs peak (typically comprising a small percentage of the total analysis window: median 0.15; inter-quartile range 0.9%) were excluded from this averaging process. Across-cycle variance (Figure 5E) was determined by the same procedure.

Vector strength of eAPs was determined as described in Goldberg and Brown (1969); its significance was evaluated with the Rayleigh test (Mardia and Jupp, 2000; $p < 0.001$). All reported correlations are Pearson's linear correlation values.

Characteristic frequency (CF) was estimated by determining peak spike rate in the responses to the binaural beats when varying carrier frequency, using the lowest SPL tested that gave a significant response (30 dB SPL, $n = 4$; 40 dB SPL, $n = 3$; 50 dB SPL, $n = 6$; 60 dB SPL, $n = 4$; 70 dB SPL, $n = 2$). The location of the peak on the rate-frequency graph was estimated by fitting a parabola to the five data points nearest to the peak value. Low tail responses to high-CF neurons can be immediately excluded for all cells (17 of 19) that responded to 500 Hz tones at 60 dB SPL or less (Figure 5 of Ohlemiller and Echteler, 1990).

Instantaneous firing rate (Figure 7C) was computed as follows. The beat-cycle average $V_B(t)$ subthreshold input potential was determined as described above, with t the time from the start of the 250-ms beat cycle. A histogram $H_0(V_B)$ of all $V_B(t)$ values was compiled using 15 equally spaced bins. The spike arrival times τ_k from the same recording were determined as described above, and the values of the cycle-averaged subthreshold potential at the corresponding times within the beat cycle , $V_B(\tau_k)$, were used to compile a histogram $H_{AP}(V_B)$, using the same binning. The conditional instantaneous firing rate p (V) was computed by

$$p(V)\mathrm{d}t = H_{AP}(V)/H_0(V),$$

with dt the sample period of $V_B(t)$. The same relation was used to predict monaural spike rates from binaural instantaneous firing rates, this time combining the p (V) derived from the binaural recording with the amplitude statistics $H_0(V)$ derived from the monaural recording. In addition to the general acceptance criteria for analysis given above, the compilation of population statistics demanded specific criteria determined by the availability of recordings suited for a particular analysis. The two cases are listed below.

1. *Binaural tuning required.* The criterion for this set of recordings was a Rayleigh test significance ($p < 0.001$) of APs to the 4-Hz beat frequency, yielding $n = 296$ beat responses, 22 cells, 3 of which were recorded in whole-cell configuration.
2. *Monaural recordings from both ears required*. This criterion yielded $n = 550$ binaural beat responses, 18 cells, 2 of which were recorded in whole-cell configuration.

Results are reported as mean ± standard deviation.