Neural Sensitivity to Interaural Envelope Delays in the Inferior Colliculus of the Guinea Pig

Sarah J. Griffin,1,2 Leslie R. Bernstein,3 Neil J. Ingham,1 and David McAlpine1,2
1Department of Physiology and 2The Ear Institute at University College London, London, United Kingdom; and 3Departments of Neuroscience and Surgery, University of Connecticut Health Center, Farmington, Connecticut

Submitted 4 August 2004; accepted in final form 27 January 2005

Griffin, Sarah J., Leslie R. Bernstein, Neil J. Ingham, and David McAlpine. Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. J Neurophysiol 93: 3463–3478, 2005. First published February 9, 2005; doi:10.1152/jn.00794.2004. Interaural time differences (ITDs) are important cues for mammalian sound localization. At high frequencies, sensitivity to ITDs, which are conveyed by the envelopes of the waveforms, has been shown to be comparable to low-frequency fine-structure–based ITD sensitivity. These brain stem pathways are not wholly independent; e.g., both involve glycinergic inhibition from the medial nucleus of the trapezoid body (Brand et al. 2002; Smith et al. 1998). Physiological recordings suggest that LSO neurons do not exclusively encode IIDs and MSO neurons do not exclusively encode ITDs. Rather, some overlap occurs (Batra et al. 1997a; Caird and Klinke 1983; Joris and Yin 1995). The strict dichotomy suggested by the duplex theory is also contradicted by psychophysical experiments that demonstrate sensitivity to ITDs conveyed by the envelopes of high-frequency complex sounds (David et al. 1959; Henning 1974; Klumpp and Eady 1956; McFadden and Pasanen 1976; Yost et al. 1971).

Human sensitivity to changes in ITD conveyed by high-frequency stimuli has typically been found to be poorer than that measured with low-frequency stimuli (Bernstein and Trahiotis 1982; Blauert 1997; Jones and Williams 1981; Yost et al. 1971). In addition, the lateral extent of intracranial images produced by ITDs conveyed by high-frequency, complex stimuli is typically smaller than that conveyed by low-frequency stimuli (Bernstein and Trahiotis 1985). The relatively poor sensitivity to ITDs at high frequencies may be explained by a lack of specialization for the processing of envelope ITDs within central auditory centers. In the LSO, ITD sensitivity in response to high-frequency stimuli has been considered a by-product of neural circuitry specialized for IID sensitivity (Joris and Yin 1995, 1998; Tollin 2003), whereas the MSO is viewed as being specialized for detection of ITDs in low-frequency stimuli.

Colburn and Equissaud (1976) hypothesized that the differences in ITD sensitivity at low and high frequencies observed psychophysically could be accounted for by differential effects of peripheral processing on low- and high-frequency stimuli. Firing of action potentials in auditory nerve fibers (ANFs) can be modeled by band-pass filtering, half-wave rectification, and low-pass filtering of the sound waveform (Geisler 1998). This produces a “phase-locked” response to low-frequency tones (Fig. 1A). When a high-frequency sound is modulated in amplitude, ANF responses exhibit phase locking to the envelope (Johnson 1980; Joris and Yin 1992; Palmer 1982; Palmer and Russell 1986). Sinusoidally amplitude modulated (SAM) tones produce firing patterns in ANFs that follow the sinusoidal envelope (Fig. 1B). Although this timing information supports sensitivity to ITDs, the firing pattern lacks the distinct “OFF periods” (where the firing probability is zero), produced by half-wave rectification, that characterize responses to low-frequency tones (compare Fig. 1, A and B).

Address for reprint requests and other correspondence: D. McAlpine, Department of Physiology, University College London, Gower Street, London, WC1E 6BT, UK (E-mail: d.mcalpine@ucl.ac.uk).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
In human psychophysical experiments, Bernstein and Trahiotis (2002) demonstrated, using “transposed” tones, that ITD sensitivity at high frequencies could be comparable to ITD sensitivity at low frequencies. Transposed stimuli were first designed by van de Par and Kohlrausch (1997) to provide high-frequency ANFs with envelope-based information similar to the waveform-based information available from low-frequency tones. The presumption is that, despite the differences in the sound pressure waveforms, the firing probability in ANFs is similar for high-frequency transposed tones and low-frequency tones (Fig. 1C).

The current study investigated neural responses to ITDs conveyed by high-frequency SAM and transposed tones; extracellular recordings were obtained from single neurons in the inferior colliculus (IC) of the guinea pig. Neural responses reflected greater sensitivity to ITDs for transposed tones than for SAM tones. Consistent with the psychophysical findings of Bernstein and Trahiotis (2002), neural discrimination of ITDs within transposed tones was comparable to neural discrimination of ITDs within low-frequency tones. Also consistent with these, and other, psychophysical (Bernstein and Trahiotis 1994; McFadden and Pasanen 1976; Nuetzel and Hafter 1981) and physiological (Joris and Yin 1998) findings, ITD sensitivity in neurons was limited to low-modulation frequencies (≤250 Hz) for both SAM and transposed tones. Our results suggest that, for those low rates of modulation, the central mechanisms that mediate sensitivity to envelope-based ITDs are essentially equivalent to those that mediate sensitivity to fine-structure–based ITDs.

Parts of this work were previously published in abstract form.

**METHODS**

**Animals and surgery**

All experiments were carried out in accordance with the Animal (Scientific Procedures) Act of 1986 of Great Britain and Northern Ireland. Single-neuron recordings were made from the central nucleus of the right IC of 27 adult guinea pigs under urethane anesthesia [Sigma-Aldrich, Poole, UK; 25% solution of 0.9% NaCl; 1 g/kg, administered intraperitoneally (ip)]. Additional analgesia was provided using fentanyl–fluanisone (Hypnorm; Janssen-Cilag, High Wycombe, UK; 0.1 ml, administered intramuscularly), supplementary doses of which were administered as required. Atropine sulfate [Animalcare, York, UK; 0.6 mg/ml; 0.1 ml, administered subcutaneously (sc)] was dispensed to reduce bronchial secretions, and lidocaine hydrochloride (Martinadale Pharmaceuticals, Romford, UK; 2%, sc) was administered locally before any surgical incision. A tracheal cannula was inserted and core temperature was maintained at 37°C with a heating blanket and rectal probe (Harvard Apparatus, Kent, UK).

Animals were placed in a sound-attenuating chamber (IAC, Winchester, UK) and held in a stereotaxic frame with hollow ear speculae (modified from model 1730, David Kopf Instruments, Tujunga, CA). Before positioning the animal, the tragus was cut to obtain clear access to the tympanic membrane. A craniotomy was performed to expose the cortex overlying the IC, the covering dura was removed, and agar (about 2%) was applied to prevent drying and deterioration of the cortex. The bullae were vented to equalize air pressure in the middle ears, by insertion of cannulae and sealing with petroleum jelly. A parylene-coated tungsten microelectrode [1–5 MΩ; World Precision Instruments, Sarasota, FL; or made in house (Bullock et al. 1988)] was positioned stereotaxically 2 mm above the IC (Medvedeva 1977) and advanced ventrally from outside the recording chamber using a piezo-stepped microdrive (Burleigh Instruments, Westbury, NY). At the end of each experiment, animals were administered a lethal dose of sodium pentobarbitone (Pentoject; Animalcare; 60 mg/ml; 1–2 ml, ip or Sagatal, 60 mg/ml).

**Stimulus production and presentation**

Sounds were produced using Tucker Davis Technologies (TDT, Alachua, FL) digital signal processing hardware. Tonal stimuli used for the isolation and characterization of single units were generated using custom software (T. Shackleton and A. Palmer, MRC Institute of Hearing Research, Nottingham, UK; 100-kHz sampling rate) and TDT system II hardware. TDT Brainware, Real Time Processor Visual Design Studio (RPVs), and system III hardware were used to generate the SAM and transposed stimuli (50-kHz sampling rate).

Stimuli were generated and scaled such that their peak voltages were at the maximum available voltage (±10 V) of the D/A converters (DACs). The outputs were attenuated to achieve the desired level for the experiments using PA4 (system II) or PA5 (system III) modules (TDT). (In initial experiments using system III, the attenuation was achieved by scaling within the signal-generation software.) Fixed amplification [Rotel (Worthing, UK) RB971 power amplifier or a Beyerdynamic (Burgess Hill, UK) A150 Blueprint stereo-amplifier] was followed by 60 dB of “final” attenuation. Using a high-level signal followed by attenuation maximized the signal-to-noise ratio through the stimulus-generation pathway. Specifically, this technique yielded a greater signal-to-noise ratio than if amplification alone were used to achieve the desired intensity. Sounds were delivered by Beyerdynamic DT48A loudspeaker drivers, modified and fitted with a probe-tube attachment to allow insertion and sealing into the hollow ear speculae. Knowles Acoustics (Burgess Hill, UK) FG3452 microphones [attached to steel tubes placed in the ear speculae and calibrated against a Bruel & Kjær (Stevenage, UK) Type 4136 1⁄8-in. microphone] were used to measure the stimulus within a few millimeters of the tympanic membrane to ensure that the sounds delivered to each ear were well matched.

**Spike collection**

Electrical signals from the electrode were conducted by a headstage to a preamplifier (TDT Medusa RA16PA) where they were amplified and digitized at a 25-kHz sampling rate. The signal was then conducted by a fiber-optic cable to the RA16 base station, which pro-
duced fixed amplification and filtering (×1,000 gain, 300-Hz high-pass filter, 10-kHz low-pass filter, and 50-Hz notch filter).

Using system II hardware, a spike voltage–level discriminator (TDT ET1) was used to detect spikes from the background noise. Spikes were monitored on a Tektronix TD210 oscilloscope for continuity of spike characteristics (i.e., shape and amplitude) to ensure recordings were from a single neuron.

Using system III hardware, spike data passed from the RA16 base station to TDT Brainware and spikes that crossed a user-defined trigger level were counted. Spikes were then additionally sorted, in Brainware, according to spike characteristics, to ensure data were collected from a single neuron.

**Sound stimuli**

**ISOLATION AND CHARACTERIZATION STIMULI.** Binaural “search stimuli,” consisting of 50-ms presentations of diotic single tones, repeated at a rate of 5/s, were used to isolate neurons and estimate their threshold (the lowest-intensity sound required to evoke firing, determined audiovisually, by the experimenter) and characteristic frequency (CF; the frequency at which the lowest threshold was obtained). The CF and threshold were confirmed by recording a frequency-versus-level response area spectrally flanking the CF estimate (2 octaves above and 4 octaves below CF) and between 10 and 90 dB of attenuation from the maximum system output of about 110 dB SPL. If the unit remained isolated for a sufficient length of time, neurons were further characterized by presentation of 50- or 200-ms diotic, ipsilateral, and contralateral tones with frequency equal to CF. Tones were presented at 20 dB above threshold and 100 or 150 repetitions were presented.

**SAM and transposed tones.** Two parameters of the stimuli were varied: ITD (of primary interest) and frequency of modulation (f_m). SAM and transposed stimuli were constructed with a carrier frequency (f_c; the frequency at which the lowest threshold was obtained) obtained. The CF and threshold were confirmed by recording a frequency-versus-level response area spectrally flanking the CF estimate (2 octaves above and 4 octaves below CF) and between 10 and 90 dB of attenuation from the maximum system output of about 110 dB SPL. If the unit remained isolated for a sufficient length of time, neurons were further characterized by presentation of 50- or 200-ms diotic, ipsilateral, and contralateral tones with frequency equal to CF. Tones were presented at 20 dB above threshold and 100 or 150 repetitions were presented.

ITD generation

Stimuli were 500 ms in duration and were gated on and off with 2-ms cosine-squared ramps. An ITD was created by delaying the entire waveform (both fine-structure and envelope) in one channel and advancing it in the other by an equal amount. The resulting stimuli thus contained onset, ongoing, and offset ITDs. A positive ITD was created by delaying the stimulus in the ipsilateral (left) ear and advancing the stimulus in the contralateral ear. A negative ITD was created by delaying the stimulus in the contralateral ear and advancing the stimulus in the ipsilateral ear. The values of ITD that were chosen depended on the f_m of the stimuli such that 17 steps of an interaural phase difference (IPD) were presented over ±0.5 cycle of the f_m. The IPDs applied remained constant irrespective of the f_m, ensuring that the neural tuning was measured over a complete cycle of each f_m. Data are also presented from experiments in which the stimuli were either 800 ms in duration or, alternatively, were 500 ms in duration with no onset or offset ITD. In the latter case, stimuli were gated on and off at the same time in both ears.

**Modulation frequency**

Neural responses from one neuron, to any one f_m and over all 17 IPDs, will be referred to as a single “recording.” SAM and transposed tones were both presented at each f_m. Initially, for each neuron, 6 or 7 f_m values were presented pseudo-randomly at each IPD. Modulation frequencies varied between 10 and 640 Hz (or CF/6 if this was lower) in logarithmic steps or between 10 and 650 Hz (or CF/6) in linear steps. The rationale for limiting the f_m to CF/6 is explained below. If the neuron remained well isolated, further data sets were collected at more narrowly spaced f_m. For 11 neurons, in later experiments, a smaller range of f_m was presented (between 10 and 200–250 Hz) because it was over this range of f_m that sensitivity to ITDs was most likely. At least 3 (3–6, median = 3) repetitions of each stimulus were presented. The exact f_m values presented were determined by the experimenter according to the response of the neuron under investigation.

**Spectral components and the f_m limitation**

SAM tones are characterized by 3 frequency components: the carrier (f_c) and 2 “sidebands” [(f_c + f_m) and (f_c – f_m)]. Transposed tones have additional sidebands spaced at multiples of 2 × f_m (Fig. 2). In psychophysical experiments van de Par and Kohlrausch (1997) demonstrated that the 5 central frequency components of a high-frequency transposed stimulus were sufficient to yield improved binaural performance at high frequencies. For transmission of the temporal structure of the modulation to a single ANF, the spectral components of SAM tones and the central 5 spectral components of transposed tones must fall within the ANF’s effective “filter” or spectral receptive field. As the f_m increases, sidebands are attenuated by the ANF’s filter in proportion to their spectral “distance” from the CF. This alters the temporal structure of the modulation transmitted to ANFs. Accordingly, the f_m in the present study was limited to a maximum of CF/6 to include f_m values at, and slightly above, the range over which we expected that the temporal structure would be preserved. This was judged from the shape of the frequency-versus-level response areas obtained from each neuron, and from reports of the equivalent rectangular bandwidth (ERB) of guinea pig ANFs (Evans et al. 1992). The maximum f_m was also limited to 640 or 650 Hz because this included, and slightly exceeded, the range of f_m at which sensitivity to ITDs has been observed at high frequencies (e.g., Bernstein and Trahiotis 1994; Joris and Yin 1998).

In their psychophysical experiments, Bernstein and Trahiotis (2002) limited the spectra of transposed stimuli to prevent the use of energy outside the psychophysically determined “auditory filter” (see Moore 1997) surrounding the center frequency of the signal under investigation. They low-pass filtered their half-wave rectified tones before multiplication with high-frequency carriers. It was not necessary to impose such a spectral limitation in the present study because
we investigated the response of single IC neurons. The spectral components influencing single neurons’ responses are necessarily limited by their spectral receptive fields. Spectral components distant from $f_m$ (outside the central 5 components) are unlikely to contribute any appreciable energy to a neuron with a CF equal to $f_m$ and at the intensities at which transposed tones were presented (10–30 dB above pure tone threshold).

High-resolution functions

“High-resolution” IPD functions were collected from 14 neurons. A value of $f_m$ was selected to which a given neuron exhibited a substantial modulation in its firing rate as a function of IPD (i.e., was found to be IPD sensitive). SAM and transposed tones with 101 (or 102) IPDs between $+0.5$ and $-0.5$ cycle, including an onset and offset ITD, were presented at this $f_m$. The duration of the stimulus was 500 ms and $\geq 6$ repetitions (6–13, median = 10) were presented, pseudo-randomly, of each stimulus condition.

Data analysis

Spike times were exported into Matlab 6.5 (The MathWorks, Natick, MA) for off-line analysis.

SENSITIVITY TO ITDS. Spike rates were determined by counting spikes over 2 time windows. First, spike rates were calculated by counting the number of spikes occurring during a 600-ms window from the onset of the stimulus at one ear to the offset of the stimulus at the other ear. This took account of the fact that for the lowest modulation rate used (10 Hz) with an ITD equivalent to 0.5 cycle of interaural phase difference, the 500-ms stimulus was gated on at one ear 50 ms before the stimulus was gated on at the other ear, and gated off 50 ms before the stimulus. Thus spikes counted over this time period included the onset and offset response to the stimuli, which could include responses to monaural stimulation at large values of ITD. Second, spikes were also counted over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation periods within the 350 ms. This middle time period was used to exclude neural responses to the onset and offset of the stimulus and to eliminate periods of monaural stimulation at large values of ITD.

In a small number of early experiments the stimulus duration was either 800 or 500 ms but IPDs were generated with no onset ITD. Recordings from these experiments (from 13 neurons) were analyzed over the middle 350 ms of the stimulus and are included in the current study.

Tuning to ITDs was calculated from the mean spike rates at each $f_m$ and over $\pm 0.5$ cycle of IPD. The following calculations were performed for each $f_m$ (i.e., for each recording) in response to SAM and transposed tones. The mean phase vector was calculated using the method of Goldberg and Brown (1969). At every $f_m$, there were 2 measurements for an IPD of 0.5 cycle (at +0.5 and −0.5 cycle) and the mean of these was used in the calculation. The mean phase vector was termed the “best” phase (BP); this was considered the IPD to which the recording was tuned. The normalized length of the vector (the vector strength, which varied between 0 and 1) was used to calculate the Rayleigh coefficient (Kuwada et al. 1987). The Rayleigh coefficient was the test statistic for a chi-squared distribution with 2 degrees of freedom (Mardia and Jupp 1999). An individual recording was classified as “ITD sensitive” if the Rayleigh coefficient was $>13.815, P < 0.01$ (Mardia and Jupp 1999; Yin and Kuwada 1983a), indicating that firing rates were not evenly distributed around a cycle of IPD. Individual neurons were classified as ITD sensitive if one or more recordings were classified as ITD sensitive.

HIGH-RESOLUTION FUNCTIONS. High-resolution IPD functions were obtained to calculate neural discrimination thresholds. ROC (receiver operating characteristic) analysis was used to estimate the smallest ITD that could be discriminated from 0 ITD, using the firing rates of a single neuron. IPDs were converted into ITDs with reference to the modulation frequency. ROC analysis has been used previously (Shackleton et al. 2003; Skottun et al. 2001) to compare neural ITD sensitivity to human psychophysical performance (see Shackleton et al. 2003 for a detailed description of the method.) For this analysis, spike rates were calculated over the middle 350-ms analysis period. To perform ROC analysis, a neurometric function was calculated for each high-resolution IPD function. The distribution of firing rates obtained at 0 ITD was compared with the distribution of firing rates at every IPD presented. The spike rates recorded from all repetitions (6–10) of each ITD form a sample distribution of the spike rates at that ITD. If 101 ITDs were presented, 101 pairs of spike rates are compared, including the comparison of the spike rates at 0 ITD with itself. The neurometric functions describe the probability of randomly selecting a spike rate from each pair of distributions and finding the spike rate at 0 ITD to be lowest.

Neurometric functions were smoothed by averaging over 3 consecutive ITDs. ITDs were considered discriminable from zero ITD if the probability (from the smoothed function) was $>0.75$ or $<0.25$ (giving 2 discrimination thresholds). We defined the just noticeable difference (JND) in ITD as the smallest ITD that is discriminable from zero ITD. No interpolation was carried out to estimate the point where the neurometric functions crossed 0.75 or 0.25; thus JNDs were, if anything, slightly underestimated. For 5 high-resolution functions, 102 (instead of 101) values of IPD were presented. In these cases no response to zero ITD was collected, and discrimination was then calculated from the closest negative ITD to zero.

PERIOD HISTOGRAMS. Entrainment of spike times to the modulation period for SAM and transposed stimuli was examined by binning spike times over one cycle of the modulation rate. Spike times over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation periods within the 350 ms, were analyzed. The cycles of the modulation period at which spikes occurred were adjusted such that 0 cycle was always with reference to the onset of sound in the contralateral ear. To quantify the degree of phase locking to the stimulus period, at each IPD and $f_m$, the vector strength was calculated. Phase locking was considered significant if the Rayleigh coefficient was $>13.815 (P < 0.01)$ (Mardia and Jupp 1999).

RESULTS

Sensitivity to envelope ITDs in SAM and transposed tones

Responses to SAM and transposed tones were obtained from 82 IC neurons with CFs >2 kHz (2.0–13 kHz, median = 4.5 kHz). For 55/82 neurons, peristimulus time histograms (PSTHs) were obtained to monaural and diotic CF tones. All neurons exhibited greater firing rates in response to tones at the contralateral ear than at the ipsilateral ear. For 70% of the sample (39/55), ipsilateral stimulation resulted in inhibition, exhibited as either a reduction in the spontaneous firing rate in response to ipsilateral stimulation or as a reduction in the rate of firing in response to a contralateral sound when stimulation was binaural.

Sixty-nine neurons were investigated using SAM and transposed tones of 500-ms duration with ITDs generated by delaying the entire waveform. Over both analysis windows (350 and 600 ms), more recordings (the response of a single neuron at one $f_m$) showed sensitivity to ITDs within transposed tones than in SAM tones. Correspondingly, more neurons were sensitive to ITDs within transposed than in SAM tones. Responses from a further 13 neurons were recorded using stimuli.
that were 800 ms in duration or had no onset ITDs. The responses of these 13 neurons were analyzed over the middle 350 ms of the stimulus duration and appeared qualitatively similar to responses recorded from neurons using stimuli of 500 ms (with onset ITDs). They are therefore included in all subsequent analyses examining responses to the middle 350 ms of the stimulus. The first column in Table 1 shows the number of neurons from which recordings were made, as well as the total number of recordings, using analysis windows of 600 and 350 ms. The second column in Table 1 shows the number of neurons and recordings in which ITD sensitivity was observed for transposed, but not SAM, stimuli. The third column shows the number of neurons and recordings in which ITD sensitivity was observed for both transposed and SAM stimuli, and the fourth column shows the number of neurons and recordings in which ITD sensitivity was observed for SAM stimuli only. The final column shows the number of neurons and recordings in which ITD sensitivity was not observed for either transposed or SAM stimuli.

Although the entire SAM and transposed waveforms were delayed, ITD sensitivity was conveyed by the envelope structure of the stimulus; neurons are not sensitive to carrier-based ITDs at frequencies >2 kHz (Joris 2003). In addition, the modulation in firing rates occurred over a cycle of interaural phase with respect to the modulation frequency and not the carrier frequency.

For the 69 neurons from which recordings were made using stimuli of 500-ms duration, more recordings were classified as ITD sensitive when spikes were analyzed over the full (600-ms) analysis window than the middle 350-ms analysis window. Recordings were considered ITD sensitive if the Rayleigh coefficient was >13.815 (see METHODS). However, the vector strengths of the recordings were, on average, lower over the 600-ms analysis window than over the 350-ms analysis window (Wilcoxon rank-sum test; \( p < 0.001 \) in response to both SAM and transposed tones). Because the Rayleigh coefficient increases as the vector strength and/or the spike count increases, the likely reason for the increased number of neurons classified as ITD sensitive when the full 600-ms window was analyzed is a result of an increase in the number of spikes counted. This appears to compensate for the reduced vector strength over the 600-ms analysis window than that over the 350-ms analysis window. All subsequent analyses pertain to responses from all 82 neurons, analyzed over the middle 350-ms window of the stimulus.

Responses of a typical IC neuron sensitive to ITDs in the envelope of SAM and transposed tones are shown in Fig. 3. The neuronal CF was 3.4 kHz, which is confirmed by the frequency-versus-level response area (Fig. 3A). PSTHs for monaural and dotic CF tones indicate the neuron to be excited by contralateral stimulation but unresponsive to ipsilateral stimulation (Fig. 3B). The response to dotic stimulation was characterized by a peak spike rate lower than that to contralateral stimulation alone, but longer in duration.

Raster plots (Fig. 3, C–H) show responses to SAM or transposed tones that were interaurally delayed over a range of ITDs equivalent to interaural phase differences \( \pm 0.5 \text{ cycle} \) of the \( f_m \) for \( f_m \) values between 10 and 510 Hz. Because the range of ITDs encompassed by \( \pm 0.5 \text{ cycle} \) of IPD differs for different modulation rates, all functions are plotted with respect to IPD. Phase locking to the envelope of the stimulus is evident for the response to both stimuli. The neuron was broadly tuned for IPD, favoring IPDs around zero. At 110 Hz \( f_m \), IPD tuning was enhanced compared with 10 Hz, with discharge rates increasing at favorable IPDs and decreasing at unfavorable IPDs. The response at favorable IPDs was greater for transposed than for SAM tones. As \( f_m \) values increased, discharge rates fell, particularly during the latter portion of the stimulus, and responses became less modulated with envelope delay. The criteria for ITD sensitivity were met in response to transposed tones at \( f_m \) values between 10 and 310 Hz and for SAM tones at 110 and 210 Hz.

Responses to SAM (left) and transposed (right) tones across all \( f_m \) and IPDs are displayed as 3-dimensional (3D) mesh plots in Fig. 3I. The neuron had a band-pass rate modulation transfer function (rMTF), with maximum firing rates at 110 Hz, a preference for IPD close to zero, and enhanced responses to transposed compared with SAM tones; peak discharge rates at the preferred \( f_m \) were almost double those evoked by SAM tones.

Responses to IPDs imposed on SAM and transposed stimuli for a second IC neuron, with a CF of 4.3 kHz, are shown in Fig. 4. PSTHs in Fig. 4B indicate that contralateral stimulation evoked higher discharge rates than binaural stimulation; ipsilateral stimulation caused discharge rates to fall below the spontaneous rate of the neuron. Raster plots (Fig. 4, C–H) show spike times for 6 recordings in response to both SAM and transposed tones from 10 to 650 Hz \( f_m \). Spike times were clearly phase locked to the envelope at 10-Hz modulation. The neuron was sensitive to IPDs in both stimuli at 10 Hz; the pattern of the phase locking to the modulation period changed with IPD and the spike rate fell around 0 IPD. At 138 Hz \( f_m \) (Fig. 4D), the neuron was IPD sensitive to transposed, but not to SAM, tones. The neuron was insensitive to IPDs in either stimulus at the higher modulation rates (Fig. 4, E–H).

Figure 4F indicates that, at 10 Hz \( f_m \), discharge rates were greater at favorable IPDs and lower at unfavorable IPDs, for transposed tones than for SAM tones. IPD sensitivity was low pass with respect to \( f_m \), unlike the band-pass response in Fig. 3I. The responses also differ from those in Fig. 3 in another important way. In Fig. 4, discharge rates were minimal for values of IPD near zero, whereas they were maximal near 0 IPD in Fig. 3. Specifically, the mean best IPD in response to transposed tones was 0.39 cycle at 138 Hz \( f_m \) (equivalent to an ITD of \( +2.826 \mu s \)), whereas, in Fig. 3, the mean best IPD was 0.03 cycle at 110 Hz \( f_m \) (equivalent to an ITD of \( +272 \mu s \)).

### Table 1. Number of neurons and recordings sensitive to ITDs, analysed over a 600- and 350-ms period

<table>
<thead>
<tr>
<th>Analysis Period</th>
<th>Total</th>
<th>Sensitive to ITDs in Transposed Tones</th>
<th>Sensitive to ITDs in Transposed and SAM Tones</th>
<th>Sensitive to ITDs in SAM Tones</th>
<th>Insensitive to ITDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 ms</td>
<td>69</td>
<td>17</td>
<td>25</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>350 ms</td>
<td>82</td>
<td>23</td>
<td>23</td>
<td>1</td>
<td>35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 ms</td>
</tr>
<tr>
<td>350 ms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of recordings</th>
</tr>
</thead>
<tbody>
<tr>
<td>600 ms</td>
</tr>
<tr>
<td>350 ms</td>
</tr>
</tbody>
</table>

\[ \text{ENVELOPE ITD SENSITIVITY} \]
Examples of recordings from 2 ITD-sensitive neurons are shown in Figs. 5 and 6. The neuron referred to in Fig. 5 (CF = 3.1 kHz) exhibited a strong onset response to diotic tones, followed by a pause and then a more sustained response. The PSTH to contralateral stimulation was similar but with a less well defined pause, whereas ipsilateral stimulation resulted in a small onset response followed by inhibition of spontaneous firing. Raster diagrams (Fig. 5, C–H) showed the response to different interaural time differences (IPDs) from 0.5 to 0.5 cycle.

Examples of recordings from 2 ITD-insensitive neurons are shown in Figs. 5 and 6. The neuron referred to in Fig. 5 (CF = 3.1 kHz) exhibited a strong onset response to diotic tones, followed by a pause and then a more sustained response (Fig. 5B). The PSTH to contralateral stimulation was similar but with a less well defined pause, whereas ipsilateral stimulation resulted in a small onset response followed by inhibition of spontaneous firing (Fig. 5B). Raster diagrams (Fig. 5, C–H) showed the response to different interaural time differences (IPDs) from 0.5 to 0.5 cycle.
and the 3D mesh plot (Fig. 5I) indicate that the neuron’s discharge rate increased with increasing $f_m$; that is, the discharge rate was high-pass with respect to $f_m$. At higher values of $f_m$, the discharge rate was greater for transposed than for SAM tones, whereas at lower values of $f_m$, the opposite was the case. Thus the discharge rate exhibited greater modulation as a function of $f_m$ in response to transposed tones. Visual inspection suggests some form of binaural interaction occurred at $f_m$ values of 10 and 20 Hz; phase locking appeared weaker near zero IPD (Fig. 5, C and D). Further, at $\leq 320$ Hz, the discharge rate was lower for values of IPD near zero, although the criteria for ITD sensitivity were not met (Fig. 5I).

**FIG. 6.** Another IC neuron (25805; CF = 2.8 kHz; threshold = $-79$ dB re maximum system output) that was not classified as ITD sensitive to either SAM or transposed tones. Figure follows the same format as Fig. 4. C–H: raster plots for recordings at 90 or 130 Hz $f_m$, which are included in I. Black bars indicating the middle 350 ms of the stimulus are not shown.
Neural responses in Fig. 6 (CF = 2.8 kHz) were also insensitive to ITDs in SAM or transposed tones. PSTHs to pure tones (Fig. 6B) indicate the neuron to be excited by contralateral stimulation and the number of spikes to be facilitated by binaural stimulation. The PSTH to ipsilateral stimulation shows inhibition of spontaneous firing, following a small onset response. In contrast to the onset response of the neuron in Fig. 3 and the onset-pause–sustained responses in Figs. 4B and 5B, this neuron had an adapting PSTH to binaural stimulation. Raster plots in Fig. 6, C–H indicate strong phase locking to the \( f_m \). At 10 Hz (Fig. 6C) it is clear that the spike times were more tightly phase locked to the \( f_m \) in response to transposed tones than to SAM tones. Although discharges were entrained to the envelopes of both SAM and transposed tones, the response to the transposed tones occurred over a shorter duration of the modulation period, reflecting the shorter “on period” (where the value of the sound pressure waveform is >0) per cycle of the transposed tone. Mean firing rates in Fig. 6I were not modulated with either \( f_m \) or IPD.

Mean firing rates of a further 9 neurons, chosen to illustrate the variety of responses obtained, are displayed as 3D mesh plots in Fig. 7. Neurons classified as sensitive to envelope-based ITDs and with firing rates that peak close to zero IPD are arranged in the top row and neurons with firing rates that peak close to IPDs of 0.5 cycle in the middle row. The bottom row shows neurons insensitive to envelope-based ITDs. For each ITD-sensitive neuron, the maximum Rayleigh coefficient (which was used as a measure of sensitivity to ITDs) was determined across all recordings in response to SAM and transposed tones, and both the top and middle rows (Fig. 7, A–C and D–F) are arranged from left to right in order of decreasing maximum Rayleigh coefficient.

All 3 neurons in the top row of Fig. 7 showed greater modulation of their discharge rate in response to transposed tones than to SAM tones, as a function of both IPD and \( f_m \). The neuron in Fig. 7A was very strongly sensitive to ITDs conveyed by transposed tones over a specific range of modulation rates (138–266 Hz), with the greatest Rayleigh coefficient at 266 Hz, whereas it was largely unresponsive to ITDs conveyed by the SAM stimulus. The other 2 neurons (Fig. 7, B and C) responded strongly to both transposed and SAM tones and were sensitive to ITDs over a specific range of \( f_m \) (about 20 and 40 Hz in Fig. 7B and about 115 and 220 Hz in 7C in response to transposed tones and about 40 Hz in Fig. 7B in response to SAM tones). Generally, when neurons were sensitive to ITDs in both stimuli, sensitivity occurred at the same \( f_m \) but extended to a wider range of \( f_m \) in response to transposed tones (e.g., in Fig. 7B). Overall, discharge rates were more highly modulated to ITDs conveyed by transposed stimuli than they were to ITDs conveyed by SAM stimuli. In addition, rates were consistently higher at favorable ITDs and lower at unfavorable ITDs for ITDs conveyed by transposed and by SAM tones. At any \( f_m \) preferred IPD tuning (where such sensitivity occurred) was similar in response to both stimuli (i.e., peaks and troughs in firing rates occurred at similar IPDs). As a function of \( f_m \), the neural response in Fig. 7C had a trough in its firing rate at 115 Hz \( f_m \), which was more pronounced (i.e., reaching a lower firing rate) in response to transposed tones than to SAM tones.

Discharge rates as a function of IPD were also more modulated in response to transposed tones for the responses in Fig. 7, D and E. Neurons were insensitive to ITDs in the envelope of SAM tones, whereas in Fig. 7, D and E, sensitivity occurred for more than one \( f_m \) in response to transposed tones.

**FIG. 7.** 3D mesh plots (same format as in Fig. 4I) for 9 IC neurons. Firing rates in response to ITDs conveyed by SAM (left) and transposed (right) tones, at different \( f_m \) values are shown. Top row: ITD-sensitive neurons with peak responses close to 0 ITD. A: neuron 25806, CF = 7.8 kHz, threshold = −44 dB re maximum system output. B: neuron 12505, CF = 3.6 kHz, threshold = −81 dB re maximum system output. C: neuron 19202, CF = 2 kHz, threshold = −52 dB re maximum system output. Middle row: ITD-sensitive neurons with a trough in response close to 0 ITD. D: neuron 27204, CF = 2.2 kHz, threshold = −55 dB re maximum system output. E: neuron 12821, CF = 8.8 kHz, threshold = −62 dB re maximum system output. F: neuron 20818, CF = 4.8 kHz, threshold = −70 dB re maximum system output. Bottom row: neurons classified as insensitive to ITDs. G: neuron 22507, CF = 5.1 kHz, threshold = −85 dB re maximum system output. H: neuron 20804, 5.4 kHz, threshold = −81 dB re maximum system output. I: neuron 25605, CF = 3.8, threshold = −73 dB re maximum system output.
Sensitivity to ITD as a function of $f_m$

The number of ITD-sensitive recordings as a proportion of the total number of recordings is displayed in Fig. 8, calculated as a function of $f_m$, in 50-Hz bins. Sensitivity to ITDs could be described as band-pass as a function of $f_m$ with the greatest proportion of ITD-sensitive recordings between 60 and 310 Hz $f_m$. The sharpest reduction in the proportion of ITD-sensitive recordings occurred as the $f_m$ increased $>$310 Hz and no recordings were considered sensitive to ITDs at $f_m$ $>$550 Hz. The total number of recordings in each $f_m$ bin is shown in the inset of Fig. 8.

The distribution of best phases (BPs) for all ITD-sensitive recordings is shown in Fig. 9. Positive BPs indicate a preference for sounds leading at the contralateral ear; negative BPs indicate a preference for sounds leading at the ipsilateral ear. BPs to both transposed tones (right) and SAM tones (left) cluster around 0 and 0.5 cycle, with more positive than negative values. This is consistent with previous physiological investigations of neural sensitivity to ITDs, which indicate that IC neurons have, predominantly, BPs corresponding to sounds leading at the contralateral ear (Caird and Klinke 1987; McAlpine et al. 2001; Yin and Kuwada 1983b; Yin et al. 1984, 1986).

For some ITD-sensitive neurons in this study it was possible to make recordings that were closely spaced in $f_m$, and the ITD tuning of these neurons as a function of $f_m$ was investigated. Responses from the 8 neurons with the highest number of ITD-sensitive recordings in response to transposed tones are shown in Fig. 10A. Discharge rates were normalized to maximum at each $f_m$ to emphasize the ITD tuning. Responses to SAM tones are shown on the left and responses to transposed tones are shown on the right. Note that abscissae are scaled to take account of the range of modulation rates at which sensitivity to ITDs was observed in each neuron. Neurons are organized by eye, according to whether maximum discharge rates were aligned around a common ITD for different $f_m$ values (top 2 rows) or whether minimum discharge rates were aligned around a common value of ITD for different $f_m$ values (bottom row). For most neurons, the positions of response maxima and response minima as a function of $f_m$ were similar for SAM and transposed tones.

Neurons in which response maxima are aligned are referred to as “peak” neurons and neurons in which response minima are aligned are referred to as “trough” neurons (Batra et al. 1997b; Yin and Kuwada 1983b). Phase plots (BP plotted as a function of $f_m$) are shown in Fig. 10B for the same neurons for which responses are shown in Fig. 10A. Theoretically, for neurons described as “peak” or “trough” neurons, BP changes systematically with pure-tone frequency or modulation frequency (Batra et al. 1997b; Joris 1996; Rose et al. 1966). The number of ITD-sensitive recordings as a proportion of the total number of recordings is displayed in Fig. 8, calculated as a function of $f_m$, in 50-Hz bins. Sensitivity to ITDs could be described as band-pass as a function of $f_m$ with the greatest proportion of ITD-sensitive recordings between 60 and 310 Hz $f_m$. The sharpest reduction in the proportion of ITD-sensitive recordings occurred as the $f_m$ increased $>$310 Hz and no recordings were considered sensitive to ITDs at $f_m$ $>$550 Hz. The total number of recordings in each $f_m$ bin is shown in the inset of Fig. 8.

The distribution of best phases (BPs) for all ITD-sensitive recordings is shown in Fig. 9. Positive BPs indicate a preference for sounds leading at the contralateral ear; negative BPs indicate a preference for sounds leading at the ipsilateral ear. BPs to both transposed tones (right) and SAM tones (left) cluster around 0 and 0.5 cycle, with more positive than negative values. This is consistent with previous physiological investigations of neural sensitivity to ITDs, which indicate that IC neurons have, predominantly, BPs corresponding to sounds leading at the contralateral ear (Caird and Klinke 1987; McAlpine et al. 2001; Yin and Kuwada 1983b; Yin et al. 1984, 1986).

For some ITD-sensitive neurons in this study it was possible to make recordings that were closely spaced in $f_m$, and the ITD tuning of these neurons as a function of $f_m$ was investigated. Responses from the 8 neurons with the highest number of ITD-sensitive recordings in response to transposed tones are shown in Fig. 10A. Discharge rates were normalized to maximum at each $f_m$ to emphasize the ITD tuning. Responses to SAM tones are shown on the left and responses to transposed tones are shown on the right. Note that abscissae are scaled to take account of the range of modulation rates at which sensitivity to ITDs was observed in each neuron. Neurons are organized by eye, according to whether maximum discharge rates were aligned around a common ITD for different $f_m$ values (top 2 rows) or whether minimum discharge rates were aligned around a common value of ITD for different $f_m$ values (bottom row). For most neurons, the positions of response maxima and response minima as a function of $f_m$ were similar for SAM and transposed tones.

Neurons in which response maxima are aligned are referred to as “peak” neurons and neurons in which response minima are aligned are referred to as “trough” neurons (Batra et al. 1997b; Yin and Kuwada 1983b). Phase plots (BP plotted as a function of $f_m$) are shown in Fig. 10B for the same neurons for which responses are shown in Fig. 10A. Theoretically, for neurons described as “peak” or “trough” neurons, BP changes systematically with pure-tone frequency or modulation frequency (Batra et al. 1997b; Joris 1996; Rose et al. 1966). The number of ITD-sensitive recordings as a proportion of the total number of recordings is displayed in Fig. 8, calculated as a function of $f_m$, in 50-Hz bins. Sensitivity to ITDs could be described as band-pass as a function of $f_m$ with the greatest proportion of ITD-sensitive recordings between 60 and 310 Hz $f_m$. The sharpest reduction in the proportion of ITD-sensitive recordings occurred as the $f_m$ increased $>$310 Hz and no recordings were considered sensitive to ITDs at $f_m$ $>$550 Hz. The total number of recordings in each $f_m$ bin is shown in the inset of Fig. 8.

The distribution of best phases (BPs) for all ITD-sensitive recordings is shown in Fig. 9. Positive BPs indicate a preference for sounds leading at the contralateral ear; negative BPs indicate a preference for sounds leading at the ipsilateral ear. BPs to both transposed tones (right) and SAM tones (left) cluster around 0 and 0.5 cycle, with more positive than negative values. This is consistent with previous physiological investigations of neural sensitivity to ITDs, which indicate that IC neurons have, predominantly, BPs corresponding to sounds leading at the contralateral ear (Caird and Klinke 1987; McAlpine et al. 2001; Yin and Kuwada 1983b; Yin et al. 1984, 1986).

For some ITD-sensitive neurons in this study it was possible to make recordings that were closely spaced in $f_m$, and the ITD tuning of these neurons as a function of $f_m$ was investigated. Responses from the 8 neurons with the highest number of ITD-sensitive recordings in response to transposed tones are shown in Fig. 10A. Discharge rates were normalized to maximum at each $f_m$ to emphasize the ITD tuning. Responses to SAM tones are shown on the left and responses to transposed tones are shown on the right. Note that abscissae are scaled to take account of the range of modulation rates at which sensitivity to ITDs was observed in each neuron. Neurons are organized by eye, according to whether maximum discharge rates were aligned around a common ITD for different $f_m$ values (top 2 rows) or whether minimum discharge rates were aligned around a common value of ITD for different $f_m$ values (bottom row). For most neurons, the positions of response maxima and response minima as a function of $f_m$ were similar for SAM and transposed tones.
Note that the
in A 27204. B response to transposed tones have

Previous investigations of sensitivity to ITDs of pure tones and
Bd 1983b). For “peak” neurons, the CD is equal to the ITD at

slope of a linear regression describing the relation between BP
and frequency is referred to as the characteristic delay (CD)
and the intercept with the ordinate axis is referred to as the
characteristic delay (CD)

Neural JNDS are shown by dashed lines for SAM (blue)
and transposed (red) tones. The abscissae differ
across plots because they reflect the presentation of different
values of fn that, given the constant range of IPDs presented,
resulted in different ranges of ITDs being presented. Neurons
in Fig. 11, A, C, and E responded best to IPDs around zero
cycle, whereas those in Fig. 11, B, D, and F responded best
to IPDs around 0.5 cycle. Consistent with the responses shown
in Figs. 3, 4, and 7, discharge rates were more highly modulated
to transposed tones than to SAM tones.

The probability of correctly discriminating between zero
ITD and all other ITDs presented is shown in the neurometric
functions in each panel of Fig. 11, A–F (right panels).
Thresholds for discrimination (see METHODS) are marked by black
horizontal lines. From these ITDs, the single value closest to
zero was taken to be a measure of the just noticeable difference
(JND). Neural JNDs are shown by dashed lines for SAM (blue)
and transposed (red) tones.

Figure 11A depicts data obtained with fn of 100 Hz. At that
rate of modulation, the neural JNDS were 644 μs for SAM and
258 μs for transposed tones. Figure 11C shows data obtained
from the same neuron in Fig. 11A, but for fn of 140 Hz. For that
rate of modulation, JNDS were 672 and 177 μs for SAM and
transposed tones, respectively. Consistent with the greater
modulation in firing rates in response to transposed tones as a
function of ITD, JNDS derived from these functions were
correspondingly lower for transposed than for SAM tones. This
is also the case in Fig. 11, B, D, and E. For the responses shown
in Fig. 11E, the neural JNDS were large, being 10 and 3.3 ms
in response to SAM and transposed tones, respectively. These
large values were, at least in part, a result of the large vari-
ability observed in the measures of firing rate compared with
the mean firing rate change over a cycle of IPD. For the
responses depicted in Fig. 11F, the JND in response to trans-

Detection thresholds for ITDs in the envelope of transposed
tones: comparison with SAM tones and low-frequency
pure tones

Of primary interest for this investigation was the neural
sensitivity to ITDs in SAM as compared with transposed tones.
To quantify the sensitivity to ITDs in response to transposed
tones and SAM tones, high-resolution IPD functions were
obtained and neural discrimination thresholds were calculated
using ROC analysis (see METHODS).

High-resolution IPD functions were obtained in response to
SAM and transposed tones from 14 neurons. For 12 neurons,
functions were obtained only at a single fn. For 2 neurons,
functions were obtained at 2 fn values (for one of these 2
neurons, a second high-resolution function was obtained in
response to transposed, but not SAM, tones). This gave a total
of 16 functions obtained in response to transposed tones and 15
functions obtained in response to SAM tones. Six examples of
such functions are shown in Fig. 11, A–F. On the left, the mean
discharge rates and SDs (error bars) are shown in response to
SAM (blue) and transposed (red) tones. The abscissae differ
across plots because they reflect the presentation of different
values of fn that, given the constant range of IPDs presented,
resulted in different ranges of ITDs being presented. Neurons
in Fig. 11, A, C, and E responded best to IPDs around zero
cycle, whereas those in Fig. 11, B, D, and F responded best
to IPDs around 0.5 cycle. Consistent with the responses shown
in Figs. 3, 4, and 7, discharge rates were more highly modulated
to transposed tones than to SAM tones.

FIG. 10. ITD tuning as a function of modulation frequency. A: normalized
firing rates, at different fn plotted as a function of ITD for 8 IC neurons (a–h)
in response to transposed (left of each pair) and SAM (right of each pair) tones:
a: 19202; b: 22207; c: 18808; d: 20709; e: 22205; f: 20818; g: 20810; and h: 27204. B: BPs plotted as a function of modulation rate for the same 8 neurons
in A in response to transposed (closed circles) and SAM (open circles) tones.
Note that the y-axes in a–d and in e–h are different.
FIG. 11. Derivation of neural ITD discrimination thresholds [just noticeable differences (JNDs)]. A–F: Left: mean firing rates and SDs in response to SAM (blue) and transposed (red) tones plotted as a function of ITD from 6 “high-resolution” functions (see METHODS). Right: 3-point averaged neurometric functions associated with the firing rates shown to the left. y-axis shows the probability of correctly discriminating the ITD of a transposed or SAM tone from zero ITD. Black lines indicate the criterion level for discrimination from zero. All discriminable ITDs are marked by a filled circle for SAM tones (blue) and transposed tones (red). Smallest discriminable ITD is marked by the dotted blue line (SAM) and dotted red line (transposed). A: neuron 22508, \( f_m = 100 \) Hz, B: neuron 20810, \( f_m = 180 \) Hz, C: neuron 22508, \( f_m = 180 \) Hz, E: neuron 22508, \( f_m = 140 \) Hz, D: neuron 27204, \( f_m = 180 \) Hz, E: neuron 23008, \( f_m = 40 \) Hz, F: neuron 25916, \( f_m = 120 \) Hz. G: JNDs in response to SAM (blue), transposed (red), and low-frequency tones (black) (from Shackleton and Palmer 2004). Median JNDs in response to SAM tones were substantially higher. For \( f_m \) > 280 Hz, the JND for that same neuron was 320 \( \mu \)s in response to transposed tones, lower than either JND at 200 Hz.

The lowest JNDs in response to transposed tones were comparable to JNDs obtained in response to low-frequency tones, at corresponding frequencies, whereas the lowest JNDs in response to SAM tones were substantially higher. For \( f_m \) > 280 Hz, no neural JNDs could be calculated. As previously discussed (see Fig. 8), responses indicating sensitivity to changes in ITD were rarely recorded at these rates and when such responses were observed they were often weakly modulated with variations in ITD. In contrast, neural JNDs in response to low-frequency tones decrease as the tone frequency increases up to nearly 500 Hz (Shackleton and Palmer 2004). JNDs in response to low-frequency tones between 280 and 400 Hz contribute to the low median JND for low-frequency tones.

Phase locking to the envelope modulation

Transposed tones are designed to evoke a similar temporal pattern of action potentials from high-frequency ANFs to that evoked in low-frequency fibers by low-frequency tones. As previously discussed, high-frequency ANFs would not be expected to show this pattern in response to SAM tones. It is thus predicted that ANFs should show “tighter” phase locking to the period of a transposed than a SAM waveform (compare Fig. 1,
We examined phase locking to the modulation rate in our sample of IC neurons. Period histograms of the number of spikes occurring throughout the period of SAM and transposed waveforms were calculated for the 4 example neurons in Figs. 3–6 and are displayed in Fig. 12, A–D, respectively. Paired responses to SAM (top) and transposed (bottom) tones are shown in response to each modulation rate, with the response to 10-Hz modulation on the left. The response is clearly phase locked to 10 Hz in Fig. 12A. In response to transposed tones, spikes occur over a narrower range of the stimulus cycle than they do in response to SAM tones. In the second pair of histograms, measured at 110-Hz modulation, phase locking occurred in response to both stimuli but was tighter in response to transposed tones. As the modulation rate increased, phase locking became weaker and was no longer apparent at 310 Hz. A similar pattern is observed in Fig. 12, B, C, and D; the phase locking is limited to the lower-modulation rates in response to both stimuli but in response to transposed tones is “tighter” than to SAM tones. This likely reflects the shorter “on period” per cycle of the transposed tones.

To summarize phase locking to the modulation period for all neurons in our sample, the vector strength was calculated at each value of $f_m$ and IPD that yielded significantly phase-locked responses (Rayleigh coefficient $>13.815, P < 0.01$). A 3-factor ANOVA for stimulus type, $f_m$, and IPD revealed each factor had a significant effect on the phase locking to the modulation rate (all $P < 0.001$). Changes in phase locking, as a function of IPD, can be seen in the raster plot in Fig. 4C.

To examine the effect of stimulus type and $f_m$ on the vector strength, the vector strength values (for all recordings that showed significant phase locking) were grouped into logarithmically spaced $f_m$ bins. The means for each $f_m$ bin are shown in Fig. 13. Overall, phase locking was greater in response to transposed than to SAM tones ($t$-test, $P < 0.001$) and was low as a function of $f_m$ for both stimuli. The vector strengths in response to both stimuli decreased from similar $f_m$ values (about 100 Hz) and the characteristics of the reduction were qualitatively similar. Phase locking is essential for coding temporal information that mediates sensitivity to ITDs in binaural neurons. The reduction in phase locking with increasing $f_m$ might underlie the limit on sensitivity to envelope-based ITDs observed at around 300 Hz (see Figs. 8 and 10) (see DISCUSSION).

DISCUSSION

The principal finding of this study is that neural sensitivity to ITDs conveyed by the envelopes of high-frequency sounds can be as great as that observed for ITDs conveyed by the fine...
structure of low-frequency sounds. The neural data presented here, recorded from the IC of the guinea pig, are consistent with the enhanced ITD sensitivity yielded by transposed tones observed in psychophysical data with human listeners (Bernstein and Trahiotis 2002). More neurons were sensitive to ITDs within the envelope of transposed than of SAM tones. Firing rates were more modulated as a function of ITD, and JNDs for ITD were consistently lower, in response to transposed tones than to SAM tones. We conclude that neurons are more sensitive to ITDs within transposed tones than within SAM tones.

Comparison with human psychophysics using transposed tones

The current study of neural responses to ITDs was motivated by the finding that, for human listeners, transposed tones confer an improvement in sensitivity to ITDs at high carrier frequencies compared with SAM tones (Bernstein and Trahiotis 2002). The neural JNDS we observed follow a pattern similar to that obtained from human listeners; threshold ITDs obtained with transposed tones were smaller than those obtained with SAM tones. It is noteworthy that Bernstein and Trahiotis (2002) found threshold ITDs to be immeasurable with the high-tones. It is noteworthy that Bernstein and Trahiotis (2002) transposed tones were smaller than those obtained with SAM tones. The neural responses to SAM and transposed tones, in the current study, were qualitatively similar to these previous studies; BPs were predominantly leading at the contralateral ear and phase plots were not necessarily well described by linear regression. Clustering of BPs close to 0 and 0.5 cycle was reminiscent of EE and EI interactions described previously in response to high- (Batra et al. 1997a) and low-frequency sounds (Yin and Kuwada 1983b).

Previous authors have reported mammalian neural sensitivity to ITDs in the envelopes of high-frequency stimuli including data using SAM tones recorded from the MSO (Batra et al. 1997a, b), LSO (Joris 1996; Joris and Yin 1995), and IC (Batra et al. 1989, 1993), noise recorded from the IC (Joris 2003), clicks recorded from the IC (Caird and Klinke 1987), and high-frequency tones with trapezoid envelopes recorded from the IC (Yin et al. 1984). The neural responses to SAM and transposed tones, in the current study, were qualitatively similar to these previous studies; BPs were predominantly leading at the contralateral ear and phase plots were not necessarily well described by linear regression. Clustering of BPs close to 0 and 0.5 cycle was reminiscent of EE and EI interactions described previously in response to high- (Batra et al. 1997a) and low-frequency sounds (Yin and Kuwada 1983b).

Comparison with previous in vivo electrophysiology

We compared neural JNDS obtained in the current study with neural JNDS obtained in response to ITDs in low-frequency tones in the IC of the guinea pig by Shackleton and colleagues (Shackleton and Palmer 2004; Shackleton et al. 2003; Skottun et al. 2001). Neural JNDS in response to transposed tones were comparable to JNDS in response to low-frequency tones at $f_m \leq 300$ Hz, consistent with human psychophysics (Bernstein and Trahiotis 2002). The current study used a longer stimulus duration (500 vs. 50 ms) and a smaller number of repetitions (6–13 vs. 100) than were used by Shackleton and colleagues. Onset and offset responses were excluded in the calculation of spike rates for ROC analysis in the current study, in contrast to the inclusion of the whole stimulus-driven response in the calculation of low-frequency JNDS by Shackleton and colleagues. The stimulus used by Shackleton and colleagues had no onset ITD. Despite the differences in stimulus characteristics, neural JNDS from transposed and low-frequency tones were similar. Finally, no qualitative differences between JNDS calculated over the 600-ms analysis window and the 350-ms analysis window were observed in response to SAM and transposed tones.

Phase locking and $f_m$ limitation on sensitivity to ITDs

We suggest that the enhanced neural ITD sensitivity observed with transposed stimuli was a result of the temporal pattern of action potential firing in high-frequency ANFs, stemming from the temporal signature of the envelopes of those stimuli. Transposed tones were originally designed to overcome the nature of peripheral auditory processing hypothesized to be responsible for the relatively poor binaural performance observed with high-frequency stimuli (van de Par and Kohlrausch 1997). The transposed stimulus was designed to produce an output at the level of high-frequency ANFs that, in terms of its temporal pattern, mimics the output normally observed at low frequencies (Fig. 1). Data describing ANF responses to transposed tones are not yet published but the analysis of phase locking in the IC presented here indicates that phase locking is enhanced in response to transposed compared with that in response to SAM tones (Fig. 13). This enhancement in phase locking does not require any increased ability to entrain to the transposed waveform but can be explained simply by the different envelopes of the stimuli. If a neuron’s pattern of firing followed exactly the time course of SAM and
transposed tones, the phase locking to the transposed waveform would be "tighter" because the transposed waveform itself occurs over a narrow range of the modulation period.

The tighter phase locking in response to transposed tones may explain the enhancement in sensitivity to ITDs observed in behavioral experiments with human listeners and in the physiological recordings presented here. Specifically, binaural neurons in the superior olivary complex (SOC) may receive temporally more precise patterns of action potentials in response to transposed tones than in response to SAM tones. Inputs arriving in phase at the SOC are more likely to produce coincident action potentials, with a high output from EE (peak) neurons, whereas inputs arriving out of phase produce fewer coincident action potentials and reduced output from EE neurons. Responses to SAM tones were still highly phase locked to the SAM waveforms and an alternative explanation for the improvement in ITD sensitivity might involve, in some direct fashion, the "OFF period" in response to transposed and low-frequency tones. Binaural inputs might not be any more likely to arrive coincidently in response to transposed tones but rather the "OFF period" might be crucial for the efficient binaural processing of ITDs. In any case, the notion that the enhancement observed with transposed tones is directly tied to the properties of the stimulus is consistent with the finding by Bernstein and Trahiotis (2002) that the enhancements observed physiologically were well accounted for by changes in the normalized interaural correlation of the waveforms as processed by the auditory periphery.

Although transposed tones are designed to provide a pattern of responses in high-frequency ANFs similar to that seen in low-frequency fibers in response to low-frequency tones, the response patterns are, in fact, unlikely to be identical. In response to low-frequency tonal stimulation auditory hair cells are hyperpolarized during the "OFF" or nonpreferred half-cycle of the motion of the basilar membrane and stereocilia are deflected toward their shortest row and the mechanically gated ion channels are closed (Fettiplace and Fuchs 1999). During these "OFF" periods, the response of the auditory nerve can fall below its spontaneous rate. Although such a reduction would not be expected in response to high-frequency transposed tones, neural JNDs for modulation rates <300 Hz are remarkably similar for transposed and low-frequency tones.

Although phase locking is greater in response to transposed tones, phase locking decreases with \( f_m \) in the same manner in response to either SAM or transposed tones (Fig. 12). This is consistent with the rate limitation of phase locking to amplitude-modulated sounds measured previously (Joris and Yin 1998; Krishna and Semple 2000), and with behavioral accounts of monaural temporal modulation transfer functions (Ewert and Dau 2000; Kohlrausch et al. 2000). Consistent with the findings of these psychophysical studies, the data from the present study revealed no clear relation between CF and the maximum \( f_m \) at which ITD-sensitive recordings could be obtained.

Joris and Yin (1998) proposed that a neural limit on sensitivity to AM stimuli and to ITDs in high-frequency sounds arises in the LSO. It is likely that any such limitation results from the membrane properties of LSO neurons [e.g., the presence or absence of ion channels capable of altering the ability of neurons to follow the temporal structure of their inputs (Barnes-Davies et al. 2004)].

**Neural JNDs and physiological detection of ITDs**

Natural sounds, such as human speech, have an envelope structure replete with onsets, offsets, and "OFF periods" (Rosen 1992). ITDs conveyed by the envelope of high-frequency sounds could be relevant to binaural tasks such as sound localization and signal detection in noise (Best et al. 2004). Would the neural JNDs measured in the current study be useful to the guinea pig? In Fig. 10G, 2 measures of the "maximum" ITD that a guinea pig might experience are marked by dotted lines. The value of \( 330 \mu s \) is the maximum ITD, at low frequencies, in homogeneous time-resolved fluorescence (HTRF) values measured using noise stimuli by Sterbing et al. (2003), whereas the value of \( 150 \mu s \) is the result of a theoretical consideration of the time sounds take to travel around a spherical head (McAlpine et al. 2001). Below 300 Hz, there is only one JND for low-frequency tones (that is \( \leq 150 \mu s \)) and below 250 Hz there are few JNDs \( \leq 330 \mu s \). Similarly, in response to transposed tones, there are few JNDs \( < 330 \mu s \). If single neurons are sufficient to encode ITDs, this suggests that the usefulness of these tuning curves is limited because the JNDs are close to the maximum ITD that can be experienced. Shackleton et al. (2003) found that the lowest JNDs in response to low-frequency tones were not generated using 0 ITD as a reference. In general, they found that JNDs were minimal with reference ITDs ipsilateral to the IC from which the neuron was

As the modulation frequency of an amplitude-modulated (AM) stimulus increases such that the spectral components (see Fig. 2) fall outside the spectral receptive field of an ANF, phase locking to the AM stimulus is lost. This forms an absolute limit on the modulation rates at which sensitivity to envelope ITDs is possible. For transposed tones, the middle 5 spectral components are required to gain a binaural advantage (van de Par and Kohlrausch 1997). Thus as \( f_m \) is increased the components that are most spectrally distant from CF are attenuated most. Thus phase locking and ITD sensitivity would be predicted to decrease. We might predict that the \( f_m \) at which ITD sensitivity can occur would increase with increasing CF because ANF spectral receptive fields widen with increasing CF (Evans et al. 1992). Bernstein and Trahiotis (1994) tested this hypothesis and found that this was not the case. They suggested, instead, the existence of some other limit on the \( f_m \) values at which ITD sensitivity occurs. To account quantitatively for their psychophysical data, Bernstein and Trahiotis (2002) included a 150-Hz low-pass filter within an interaural correlation–based model of ITD sensitivity to simulate the limitation on ITD sensitivity imposed by the modulation rate. The filter they used was similar to that used in accounts of monaural temporal modulation transfer functions (Ewert and Dau 2000; Kohlrausch et al. 2000). Consistent with the findings of these psychophysical studies, the data from the present study revealed no clear relation between CF and the maximum \( f_m \) at which ITD-sensitive recordings could be obtained.
recorded. Similarly, there may be regions of the IPD tuning curves in response to transposed tones that have greater accuracy for ITD discrimination. We do not propose that single neurons are sufficient to encode ITDs; it is more likely that the responses of many neurons are pooled.

Despite limitations in the temporal coding of neurons to high-frequency sounds, the results presented here demonstrate that under conditions in which binaural neurons receive appropriate spike patterns, sensitivity to ITDs conveyed by high-frequency stimuli can be equivalent to that observed in response to low-frequency stimuli. This suggests, as first conjectured by Colburn and Esquissaud (1976), that mechanisms underlying ITD sensitivity in low- and high-frequency channels of the auditory system are, to a first approximation, equivalent.

ACKNOWLEDGMENTS

We acknowledge T. Shackleton and A. Palmer for use of software.

Present address of N. J. Ingham: Centre for the Neural Basis of Hearing, Physiological Laboratory, Downing Street, Cambridge, CB2 3EG, UK.

GRANTS

S. J. Griffin is sponsored by the Wellcome Trust 4-year program in Neuroscience. This work was funded by a Medical Research Council Grant to D. McAlpine and by National Institute on Deafness and Other Communication Disorders Grant DC-04147 to P.I.L. Bernstein.

REFERENCES

Barnes-Davies M, Barker MC, Osmani F, and Forsythe ID. Disorders Grant DC-04147 to P.I.L. Bernstein.

REFERENCES


Jones PJ and Williams RP. An experiment to determine whether the interaural time differences used in lateralizing middle and high-frequency complex patterns is dependent in any way on fine structure information. Acustica 47: 164–169, 1981.


J Neurophysiol • VOL 93 • JUNE 2005 • www.jn.org


