

# Auditory Temporal Processing: Responses to Sinusoidally Amplitude-Modulated Tones in the Inferior Colliculus

B. SURESH KRISHNA AND MALCOLM N. SEMPLE

*Center for Neural Science, New York University, New York, New York 10003*

Received 13 July 1999; accepted in final form 28 March 2000

**Krishna, B. Suresh and Malcolm N. Semple.** Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 84: 255–273, 2000. Time-varying envelopes are a common feature of acoustic communication signals like human speech and induce a variety of percepts in human listeners. We studied the responses of 109 single neurons in the inferior colliculus (IC) of the anesthetized Mongolian gerbil to contralaterally presented sinusoidally amplitude-modulated (SAM) tones with a wide range of parameters. Modulation transfer functions (MTFs) based on average spike rate (rMTFs) showed regions of enhancement and suppression, where spike rates increased or decreased respectively as stimulus modulation depth increased. Specifically, almost all IC rMTFs could be described by some combination of a primary and a secondary region of enhancement and an intervening region of suppression, with these regions present to varying degrees in individual rMTFs. rMTF characteristics of most neurons were dependent on sound pressure level (SPL). rMTFs in most neurons with “onset” or “onset-sustained” peri-stimulus time histograms (PSTHs) in response to brief pure tones showed only a peaked primary region of enhancement. The region of suppression tended to occur in neurons with “sustained” or “pauser” PSTHs, and usually emerged at higher SPLs. The secondary region of enhancement was only found in eight neurons. The lowest modulation frequency at which the spike rate reached a clear peak (“best modulation frequency” or BMF) was measured. All but two mean BMFs lay between 0 and 100 Hz. Fifty percent of the 49 neurons tested over at least a 20-dB range of SPLs showed a BMF variation larger than 66% of their mean BMF. MTFs based on vector strength (tMTFs) showed a variety of patterns; although mostly similar to those reported from the cochlear nucleus, tMTFs of IC neurons showed higher maximum values, smaller dynamic range with depth, and a lower high-frequency limit for significant phase locking. Systematic and large increases in phase-lead commonly occurred as SPL increased. rMTFs measured at multiple carrier frequencies ( $F_c$ s) showed that the suppressive region was not the result of sideband inhibition. There was no systematic relationship between BMF and  $F_c$  of stimulation in the cells studied, even at low carrier frequencies. The results suggest various possible mechanisms that could create IC MTFs, and strongly support the idea that inhibitory inputs shape the rMTF by sharpening regions of enhancement and creating a suppressive region. The paucity of BMFs above 100 Hz argues against simple rate-coding schemes for pitch. Finally, any labeled line or topographic representation of modulation frequency is unlikely to be independent of SPL.

## INTRODUCTION

Sinusoidally amplitude-modulated (SAM) tones induce in human listeners, over different modulation frequency ranges, a

variety of percepts such as fluctuation, roughness, and pitch. SAM signals have also been used in psychophysical experiments within a linear systems framework to study the temporal filtering properties of the auditory system, as distinguished (but not separate) from its spectral filtering properties (e.g., Dau et al. 1997; Salvi et al. 1982; Viemeister 1979). It is known that human speech and animal communication sounds contain many amplitude-modulated features (e.g., Rosen 1992). In particular, the envelopes found in natural speech have been shown to contain sufficient information below 50 Hz to allow a high degree of speech recognition performance, even in the relative absence of spectral cues (Shannon et al. 1995). Further, it has been reported that training language-learning impaired children with modified speech that included enhancement of envelope components in a similar frequency range (3–30 Hz) led to significant improvements in performance on various speech and language tests (Tallal et al. 1996). Understanding the processing of modulated signals is also of relevance to the design of stimulation strategies for cochlear implants, where one attempts to convey information about speech signals with a temporally varying stimulation strength at a limited number of points on the cochlea.

There now exist considerable data on physiological responses to SAM tones in the auditory nerve (e.g., Joris and Yin 1992) and cochlear nucleus (e.g., Frisina et al. 1990; Moller 1974; Rhode 1994; Rhode and Greenberg 1994). Neurons in these lower levels of the ascending auditory pathway convey information about the envelope by “phase locking” to the modulation frequency of the SAM tone. Spike rate variations with modulation frequency are generally poorly tuned or non-existent in the auditory nerve and for most neuronal types in the cochlear nucleus. In contrast, neurons at higher levels of the auditory system, including the inferior colliculus (IC), have been shown to exhibit large variations in average spike rate as the modulation frequency of the SAM tone is varied (e.g., Langner and Schreiner 1988; Rees and Moller 1983; Schreiner and Urbas 1988). It has been suggested that this represents a transformation from a “temporal code” for modulation frequency in the auditory nerve and cochlear nucleus to a “rate code” at higher levels, and that this transformation is essentially complete at the level of the IC (Langner and Schreiner 1988). However, many details of the parametric dependence of the responses to SAM tones of single neurons at locations other

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.

Address reprint requests to M. N. Semple (E-mail: mal@cns.nyu.edu).

than the auditory nerve and cochlear nucleus remain unclear. As a result, existing models of the processing of amplitude-modulation (AM) at higher levels of the auditory system are poorly constrained by experimental data and hence difficult to evaluate.

The IC is an obligatory relay in the primary lemniscal pathway from the extensively interconnected auditory mid-brain network to the cortex. It receives ascending projections from various ipsilateral and contralateral sites in the auditory periphery and descending projections from the cortex (Oliver and Huerta 1991). As a result of this extensive set of excitatory and inhibitory inputs and because it is the primary source of projections to the thalamus and cortex, it is often considered to perform an "integrative" role in auditory processing. In particular, in contrast to auditory nerve fibers and most neurons in the cochlear nucleus, prior studies (e.g., Heil et al. 1995; Langner and Schreiner 1988) of the responses of IC neurons to SAM tones with different modulation frequencies (and all other parameters fixed) have reported that most neurons in the IC show systematic variations of average spike rate with modulation frequency, with neurons often showing a clear maximum response at a best modulation frequency (BMF). However, studies (e.g., Rees and Møller 1983; Rees and Palmer 1989) that have varied other parameters of the stimulus have indicated that response patterns to SAM tones can be strongly dependent on stimulus parameters other than modulation frequency, like the sound pressure level (SPL) of stimulation. Nonetheless, a systematic description of the spike rate and synchronization properties of single IC neurons to SAM tones with a wide range of parameters is still lacking. Such a description would help clarify the implications of the emergent rate tuning in the auditory midbrain, both for neural processing mechanisms and for the representation of modulation frequency in the auditory system.

With this in mind, we present here the results of a detailed characterization of extracellularly recorded responses from single neurons in the physiologically defined central nucleus of the IC of the Mongolian gerbil (*Meriones unguiculatus*) to SAM tones with a range of modulation frequencies presented via a closed system to the contralateral ear at multiple modulation depths, SPLs, and carrier frequencies ( $F_c$ s).

## METHODS

### *Surgical and recording procedures*

Adult Mongolian gerbils with clean external ears and no sign of middle ear infections were initially anesthetized with an intraperitoneal injection of pentobarbital sodium (60 mg/kg) and subsequently maintained in an areflexive state with supplemental intramuscular injections of ketamine hydrochloride (approximately  $30 \text{ mg} \cdot \text{kg}^{-1} \cdot \text{h}^{-1}$ ). A heating pad was used to maintain a constant rectal temperature of  $37^\circ\text{C}$ . The pinnae were removed, and craniotomy of the interparietal bone was performed to expose the cerebellum. The animal was then transferred to a double-walled sound-attenuated room (IAC), where sound delivery speculae were sealed to the temporal bone and muscle around the opening of the external auditory meatus bilaterally.

Activity of single units was recorded with platinum-plated, glass-insulated tungsten microelectrodes advanced into the IC through the intact cerebellum with a stepping motor microdrive (CalTech). Electrodes typically had exposed tip lengths of 5–10  $\mu\text{m}$  and an impedance of 2–5 M $\Omega$  at 1 kHz. Electrical signals recorded from the brain

were amplified (variable gain) and filtered (typically 0.25–10 kHz). Neural signals were displayed on an oscilloscope and fed to an audio monitor. Single units were identified and isolated using the criteria of waveform similarity and separation of successive spikes by an absolute refractory period; only very well isolated single units were studied. An event timer (MALab, Kaiser Instruments) logged the occurrence of discriminated action potentials together with stimulus zero-crossings to a resolution of 1  $\mu\text{s}$ . Event times were stored in a FIFO buffer from which they were retrieved by the host computer. Acoustic stimuli were presented as the electrode was advanced to isolate responsive units. As the electrode passed through the dorsal mantle of the IC, responses that were broadly tuned, driven better by broadband noise than tones, and often habituating to repeated tonal stimulation were found. Occasionally, a coarse descending tonotopic sequence was encountered. Entry into the physiologically defined central nucleus was always marked by an increase in spontaneous background activity and the beginning of a clear ascending tonotopic progression, with narrowly tuned tone response areas that showed no signs of habituation to repeated stimulation. Within the central nucleus, both modulated and unmodulated tonal stimuli were effective in revealing responsive units.

### *Stimulus generation and data acquisition*

Stimulus waveforms were generated by two digital synthesizers controlled by a microprocessor and custom hardware for timing, data logging, and waveform control (MALab, Kaiser Instruments). The dedicated microprocessor communicated with a host computer via an IEEE-488 interface. Stimuli were digitally attenuated and transduced by electrostatic earphones (Stax Lambda, housing designed by G. Sokolich) coupled to the ear pieces. For each individual animal, SPL (expressed in dB re 20  $\mu\text{Pa}$ ) near the tympanic membrane was calibrated for both ears from 40 Hz to 40 kHz under computer control using a previously calibrated probe tube and a condenser microphone (Bruel and Kjaer, type 4134). Phase was calibrated for both ears from 40 Hz to 5 kHz by measuring the disparity in phase between a reference sinusoid generated electrically and the recorded acoustic signal. Note that all the data in this paper result from monaural contralateral presentation of sound. The magnitude transfer function of the speaker was usually smooth, with a slow rolloff until 25 kHz, the maximum frequency used; two or three local resonances, never deviating by more than 10 dB from the baseline, were also present. Deviations of the phase response from linearity were small, and not usually present. Appropriate compensations were made to the carrier prior to modulation. All acoustic stimuli were shaped digitally with a cosine-squared ramp of 5–50 ms rise time to reduce spectral splatter at onset and offset. The SAM tone is represented by the formula  $A \sin(\omega_c t)[1 + m \sin(\omega_m t)]$ , where  $\omega_c$  is the angular frequency of the carrier,  $\omega_m$  that of the modulator,  $A$  the amplitude of the carrier,  $t$  the time after signal onset, and  $m$  the modulation depth (0–100%). Such a stimulus, if sufficiently long, has a simple three component spectrum centered at  $F_c$ , as shown in Fig. 1. The SPL of the SAM tone is defined as the SPL of the carrier that is being modulated. The modulation frequency is the frequency corresponding to  $\omega_m$ .

Frequency response areas, either at a single SPL or at multiple SPLs, were recorded using pure tones (usually 100 ms, with a rise time of 10 ms, and a repetition rate  $<2.5$  Hz) and the best frequency (BF) was noted. Following this, a complete spike rate versus SPL function [rate-level function (RLF)] was usually measured (at BF). Neurons were then studied using pure tones (at multiple SPLs) and SAM tones while monitoring the responses (spike rate and vector strength) on-line to gain an initial qualitative characterization of the neuron's response properties. This initial exploration helped confine the formal presentation to the ranges of levels, depths, carrier frequencies, durations, and modulation frequencies where clear variations in response were present, thus maximizing the use of the limited recording time available. Modulation frequencies were varied within

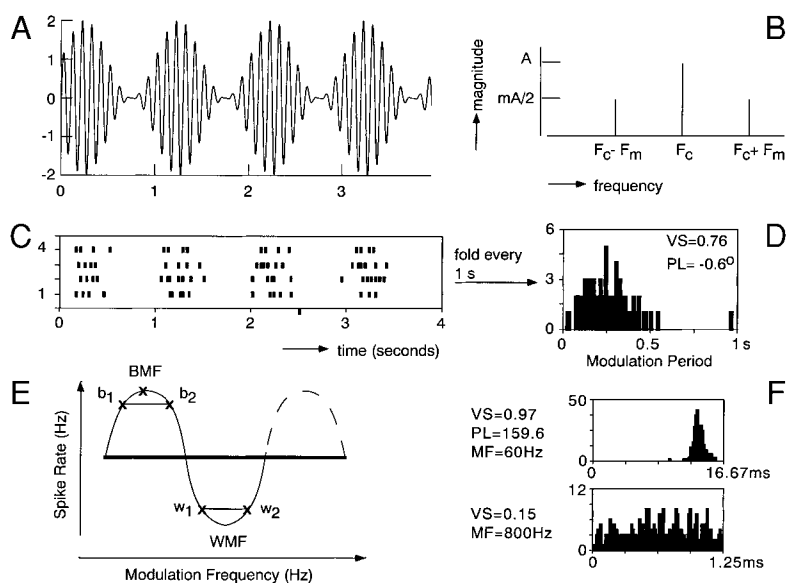


FIG. 1. Summary of stimulus characteristics and response measures. *A*: amplitude vs. time plot of a sinusoidally amplitude-modulated (SAM) tone. Carrier frequency, 10 Hz; modulation frequency, 1 Hz; modulation depth, 100%. Carrier amplitude is arbitrarily set to 1. Actual stimuli used had  $F_c$ s from 40 Hz to 25 kHz. *B*: the spectrum of an infinitely long SAM tone has 3 components.  $F_c$ , carrier frequency;  $F_m$ , modulation frequency;  $A$ , amplitude;  $m$ , modulation depth. *C*: raster plot showing cartoons of spike responses demonstrates locking to the envelope of the stimulus depicted in *A*. *D*: modulation period histogram generated from the spike trains in *C* (binwidth = 1 ms) demonstrates that spikes tend to occur only over a certain phase-range ("phase-locking"). Vector strength is 0.76 and the response shows a small negative phase-lead of  $-0.6^\circ$  relative to the stimulus envelope. *E*: the 1st peak in the rate modulation transfer function (rMTF) was chosen as the best modulation frequency (BMF; see METHODS and Fig. 18 legend). WMF, worst modulation frequency. *F*: examples of period histograms at different modulation frequencies (MF) and their associated vector strengths (VS) and phase leads (PL). Narrow histograms give vector strengths closer to 1.

the range of 0.1 Hz to a value just less than the  $F_c$  used. Stimulus durations varied from 1 to 10 s (this was determined partly by the period of the lowest modulation frequency used, so that at least 1 period was presented), while rise times varied from 5 to 50 ms. Unless it was itself the parameter being varied,  $F_c$  was set to the BF of the neuron (as measured using pure tones) at the SPL being tested. If responses were being recorded at multiple SPLs,  $F_c$  was kept constant and usually equal to the BF at an SPL in the middle of the spanned SPL range. Because it was not uncommon for the BF to vary with SPL (e.g., Kuwada et al. 1984),  $F_c$  sometimes differed slightly from the BF at some SPLs. In any case, the BF can also depend on the kind of stimulation (SAM tone vs. pure tone), the duration of stimulation (100 ms vs. 1–10 s) and the modulation frequency of stimulation (see RESULTS).

### Data analysis

Three descriptive measures of the pure-tone response are used in this paper. The first is the mean spike rate averaged over the duration of stimulation. The second is the minimum (across all SPLs tested) mean (averaged across stimulus presentations) latency at the BF of the neuron or close to it (if the BF varied with SPL). Finally, neurons were classified into one of four classes (onset, onset-sustained, pauser, and sustained) on the basis of their peri-stimulus time histograms (PSTHs) (as in Le Beau et al. 1996). However, because PSTHs can change with SPL, we also used some additional criteria: neurons were classified as pausers if they showed a pauser pattern at any SPL, and neurons were classified as sustained or onset types only if they showed this pattern at all SPLs. The few broad onset neurons found were included in the onset-sustained category, and since the identification of choppers requires more tone presentations than we used, these were not separately identified.

The response to AM is also characterized by three measures: mean spike rate (averaged over the duration of stimulation), vector strength at the modulation frequency<sup>1</sup> (Goldberg and Brown 1969) and the

mean phase-lead of the response relative to the modulating sinusoid. The functions describing the variation of each of these measures with modulation frequency will be referred to as modulation transfer functions (MTFs). MTFs plotted using the spike rate, vector strength, and mean phase-leads are called rate modulation transfer functions (rMTFs), temporal modulation transfer functions (tMTFs), and phase modulation transfer functions (pMTFs), respectively.

The vector strength is a measure of the synchrony to the modulating waveform and is equal to  $F_1/F_0$ , where  $F_1$  is the spectral magnitude of the response at  $\omega_m$  and  $F_0$  the average spike rate. It varies from a minimum of zero to a maximum of one, and as a reference, the vector strength of the sinusoidal modulating waveform of a 100% depth SAM tone is 0.5, while that of a half-wave rectified sinusoid is 0.784. A response with all spikes at precisely the same unique phase has a vector strength of 1 ("perfect" phase locking). The significance of the vector strength was assessed using the Rayleigh statistic (Stephens 1969) at the 1% significance level. Additionally, to minimize the contribution of onset spikes, at least six spikes per stimulus presentation were required for significance. The mean phase-lead was computed as  $(90 - \phi)$ , where  $\phi$  is the direction of the mean vector in the vector strength calculation. The mean phase-lead is equal to the phase of the spectral component of the response at  $\omega_m$ , relative to that of the modulating waveform. (The direction of the mean vector of the modulating waveform is  $90^\circ$ , since it produces a unimodal nonnegative period histogram symmetric about its peak at  $90^\circ$ .) The mean phase-lead was decremented by an appropriate multiple of  $360^\circ$ , i.e., unwrapped, whenever it went from a response close to  $360^\circ$  to one closer to  $0^\circ$ . pMTFs were truncated when the sampling resolution became poor enough that the phase might have skipped more than one cycle; this was not usually required, and usually only affected values above 300 Hz. Only significant vector strengths and their associated mean phase-leads are shown in the plots in this paper. Consistent with

<sup>1</sup> The measures traditionally used to characterize responses to SAM tones have been spike rate and vector strength. These are measures of the spectral magnitude of the response spike train at zero frequency and at the modulation frequency ( $F_m$ ) respectively. As a result of the nonlinear processes leading to spike generation in the auditory nerve (including the nonlinear rectification and low-pass filtering in the inner hair cell that underlies the demodulation of the SAM tone), single auditory nerve fibers respond to SAM tones with a spike train that has considerable power at these two frequencies, with the power at  $F_m$  decreasing at high frequencies (above about 1500 Hz in fibers with high

characteristic frequencies: see Joris and Yin 1992). However, in addition, the auditory nerve fiber response may also contain power at additional frequency components, like the carrier frequency ( $F_c$ ),  $F_c \pm F_m$ ,  $F_c \pm 2F_m$ , and multiples of  $F_m$  (Khanna and Teich 1989). In certain situations (e.g., Fig. 15D), it is possible that spectral magnitude at these additional frequencies may be as large as or even larger than that at  $F_m$  itself (e.g.,  $F_c - F_m$ , when it is much lower than  $F_m$ ). Further, it is possible that these components (at frequencies other than zero and  $F_m$ ) may carry stimulus-related information, in addition to that carried by the components at zero frequency (spike rate) and at  $F_m$ . A complete characterization of the temporal information contained in the IC cell's response would require a more detailed analysis of the response spike trains. However, such an analysis is beyond the scope of this paper.

the fact that phase locking to pure tones is common only below 600 Hz (Kuwada et al. 1984), we observed significant vector strength at the carrier frequency only in the few neurons in our sample with BFs in that frequency range. Synchrony to the carrier will not be discussed further in this report.

Most MTFs did not vary much when different 1-s time windows after the beginning of the stimulus were chosen for analysis, even though most neurons showed an adaptation of their mean firing rate during the stimulation period. All measures were therefore calculated over the entire stimulation period.

A BMF was defined for each rMTF as follows: first, the modulation frequency that elicited the maximum spike rate was identified. If there were two distinct maxima (e.g., Fig. 3A), the one at the lower modulation frequency (primary peak) was taken. Following this, a range of modulation frequencies ( $b_1$  to  $b_2$ , Fig. 1E) where the response was  $>90\%$  of this response maximum was extracted. The BMF was chosen as the mean of  $b_1$  and  $b_2$ . This procedure essentially corrected for skewed or irregular peaks and was almost always close to the BMF as chosen by eye. BMFs were only measured if the spike rate dropped by at least 70% on both sides of the BMF. If a 70% drop was only present on the high-frequency side, then the modulation frequency (higher than the BMF) that elicited 90% of the maximum response was chosen as the corner frequency of the rMTF. A cutoff frequency was also measured; this was the frequency at which the response fell to the minimum spike rate plus 10% of the difference between the maximum and minimum spike rates, on the high-frequency side of the primary peak ( $w_1$  in Fig. 1E). Finally, a worst modulation frequency (WMF) was extracted in rMTFs with a clear suppressive region, as evidenced by a lower spike rate in comparison with that to one or more lower depth stimuli. The method was similar to that used for the BMF: the mean of a 10% range ( $w_1$  to  $w_2$ , Fig. 1E) was taken as the WMF. In all cases, linear interpolation was used if required.

Kendall's  $\tau$  (Press et al. 1993) is used often in this paper as a nonparametric measure of correlation between two or three variables.

## RESULTS

We measured responses from 109 single neurons in the physiologically characterized central nucleus of the IC in 34 Mongolian gerbils. Three onset neurons were unresponsive to AM across the entire parameter space, a proportion similar to that reported earlier (Rees and Palmer 1989).

### Varying modulation depth

Varying SAM tone modulation depth revealed that IC rMTFs were composed of 1 or 2 regions of enhancement (modulation frequency ranges where the spike rate increases with increase in stimulus modulation depth) and/or 1 region of suppression (a range of modulation frequencies where the spike rate decreases as stimulus modulation depth increases). Figures 2 and 3 show illustrative examples of rMTFs and tMTFs measured from individual neurons (with a range of BFs) using SAM tones with varying modulation depths and constant SPL. In all cases (except Fig. 2D, see legend), the  $F_c$  was equal to the BF of the neuron at that SPL. The rMTFs in Fig. 2 show examples of neurons with rMTFs that predominantly show either an increase (A–C) or decrease (D) in spike rate with increasing modulation depth at all modulation frequencies. The magnitude of change depends on the modulation frequency. These changes create band-pass (A–C) or bandsuppressive (D) rMTFs that are characterized by their single

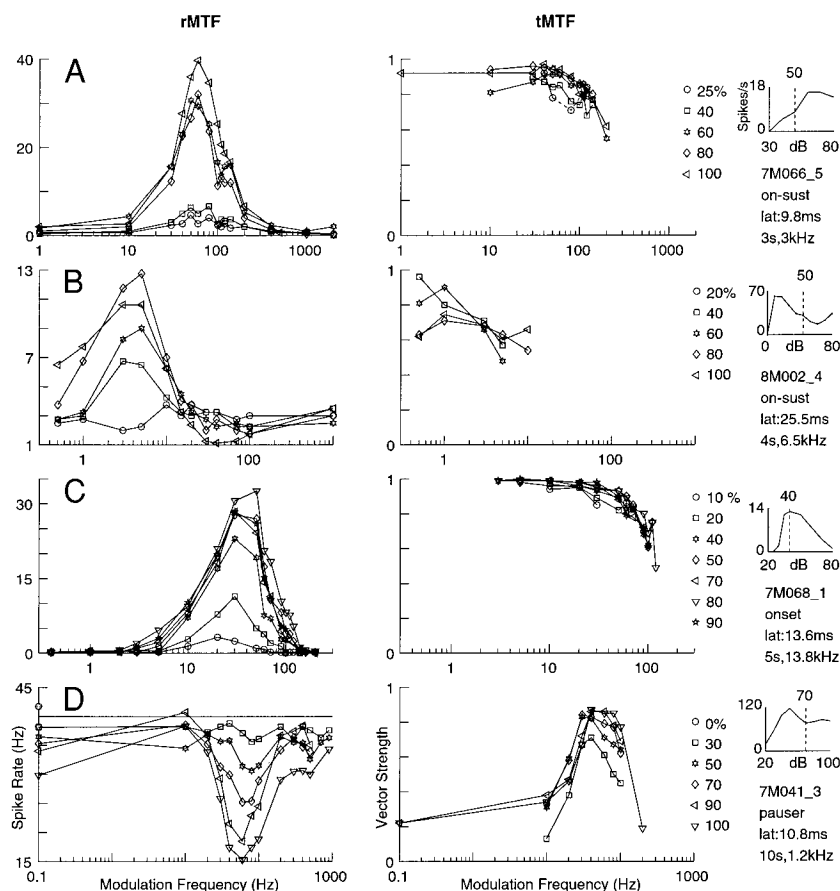


FIG. 2. Effects of varying modulation depth. Spike rate can be either enhanced (A–C) or suppressed (D) as stimulus modulation depth increases. Each row (Fig. 2, A–D) depicts responses from 1 neuron. The 1st column shows rMTFs, and the 2nd column temporal MTFs (tMTFs) for 4 neurons (A–D) studied with SAM tones of different modulation depths (identified by different symbols) and  $F_c$  at BF [for the tested sound pressure level (SPL)]; except in D, where  $F_c$  was below the BF of 1–5 kHz (see Fig. 4 for this neuron's response at BF). Dashed lines between points in tMTFs indicate that (unplotted) responses at some intervening modulation frequencies did not reach significance. In case of a zero depth stimulus (unmodulated tone), the mean of the responses (points in gray) is shown with a horizontal line (e.g., Fig. 2D). The right column shows the rate-level function (RLF), and the vertical dashed line shows the SPL at which the SAM tone was presented. The text below gives the neuron number, peri-stimulus time histogram (PSTH) class, minimum mean latency, SAM tone duration, and  $F_c$ . Tone-pip duration, rise time, number of presentations (per data point in the RLF), SAM tone rise time, and number of presentations (per data point in the MTFs) in each case were as follows, respectively. A: 100 ms, 10 ms, 10, 30 ms, 1. B: 100 ms, 10 ms, 10, 50 ms, 1. C: 100 ms, 10 ms, 10, 30 ms, 1. D: 100 ms, 10 ms, 10, 10 ms, 1.

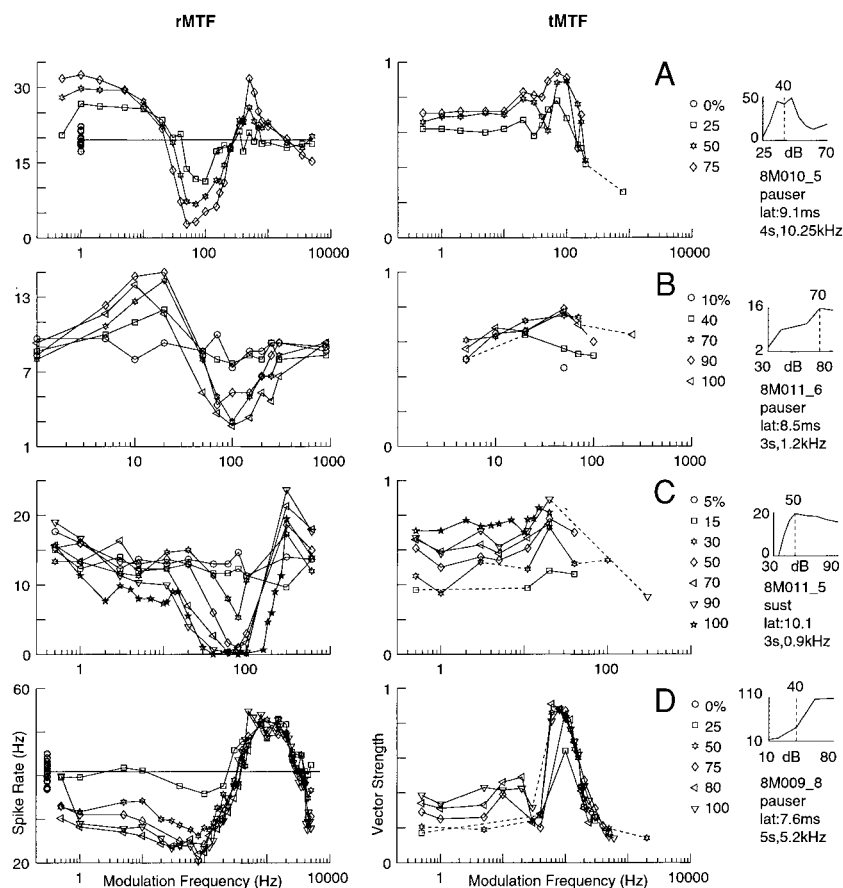


FIG. 3. Effects of varying modulation depth. A single rMTF can show both enhancement and suppression (at different modulation frequencies) as the stimulus modulation depth increases. The format of this figure is identical to that of Fig. 2. Stimulus parameters as in Fig. 2 were as follows. A: 100 ms, 10 ms, 10, 50 ms, 1. B: 300 ms, 10 ms, 10, 50 ms, 1. C: 300 ms, 10 ms, 10, 50 ms, 1. D: 250 ms, 5 ms, 15, 30 ms, 1.

prominent region of enhancement and suppression, respectively. In contrast, Fig. 3 illustrates examples of neurons whose rMTFs show both regions of enhancement and suppression. In other words, both the direction (increase or decrease) and magnitude of change (in spike rate with modulation depth) depends on the modulation frequency. The rMTFs in Fig. 3, A, C, and D, show secondary peaks at higher modulation frequencies; these were only found in eight neurons (see DISCUSSION). However, in contrast to responses at the primary peak (BMF), responses at the secondary peak possessed low vector strength; i.e., they were not synchronized to the modulation frequency. We formally recorded the effects of systematic variation of modulation depth from 31 neurons; in most cases, the spike rate at any given modulation frequency increased or decreased almost monotonically with increasing modulation depth.

tMTFs ranged from low-pass to band-pass (more peaked) shapes; this is related to the SPL of stimulation, as will be shown in a subsequent section. Notice that the vector strength remains high (and often reaches its peak) as the spike rate reaches its minimum (Figs. 2D and 3). Also, low modulation depths can elicit responses with high vector strengths, i.e., the dynamic range of the vector strength measure is small as modulation depth is varied. For example, the neuron in Fig. 2C showed a vector strength  $>0.8$  in response to a 10% stimulus modulation (an increase of 0.82 dB and a decrease of 0.91 dB from the mean SPL) at 30 Hz. In contrast to the rapid saturation of the vector strength of the response (VSr), the vector strength of the modulating waveform (VSs; equal to half the modulation depth) increases linearly with depth from 0 to 0.5. As a result, an alternative measure of synchrony, modulation gain [20

$\log(\text{VSr}/\text{VSs})$ ] (e.g., Frisina et al. 1990) decreases almost monotonically as the depth increases. No consistent trend was found in the behavior of pMTFs at different modulation depths, and in most cases, the magnitude of variation was small.

#### Varying sound pressure level

The rMTFs for each of the neurons in Figs. 2 and 3 were recorded at a fixed SPL and  $F_c$ . Therefore it is of interest to know whether the observed rMTFs are invariant characteristics of the neuron, independent of the particular parameters (SPL and  $F_c$ ) chosen for stimulation. Figure 4 partially answers this question by demonstrating that it is possible for a region of enhancement in the rMTF from a single neuron to be transformed systematically into a region of suppression as the SPL is increased. Note that these rMTFs were measured with the carrier at 1,500 Hz, while the effects of using a 1,200-Hz carrier are shown in Fig. 2D for the same neuron. The position of the rate minimum is not affected by this change in  $F_c$ , suggesting that spectral effects (like the positioning of the AM sidebands in putatively inhibitory regions of the frequency response area) are not important in its generation (see subsequent section on effects of varying carrier frequency). The tMTFs at multiple SPLs and 100% depth show a systematic “low-pass to band-pass shift,” as a result of the decrease in vector strength at low modulation frequencies with increasing SPL (Fig. 4F).

To explore the emergence of suppression at higher SPLs further, we systematically varied the SPL while keeping the depth at 100% and the carrier at or close to the BF in 56

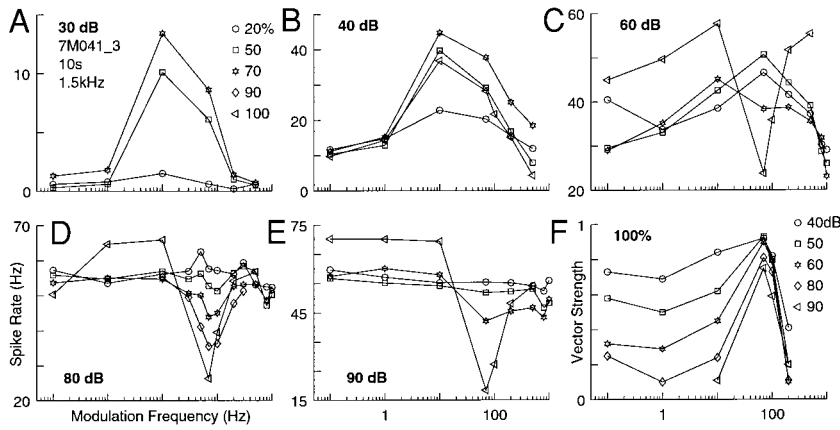


FIG. 4. The region of enhancement may become a region of suppression with increasing SPL. A–E: rMTFs measured at different SPLs (increasing from A to E) and multiple modulation depths (see key in A). The common  $F_c$  and duration of stimulation for A–F are given in A. rMTFs measured at 50, 70, and 100 dB (not shown) behaved intermediate to the rMTFs at the flanking SPLs. Note that the ordinate scale varies while the abscissa is identical in A–F. F: tMTFs at 100% depth and multiple SPLs (see key in F) are plotted for the same neuron (also see Fig. 2D). Each data point in A–F resulted from a single presentation of the SAM tone (10-ms rise time).

neurons. Illustrative examples spanning the range of behaviors seen are shown in Figs. 5 and 6. Figure 5 shows examples of MTFs at multiple SPLs from four neurons with a range of best frequencies and with rMTFs that remain band pass (but with varying bandwidths) at all tested SPLs. The neurons depicted in A and B responded at all modulation frequencies tested, but suffered a loss of tuning at higher SPLs (possibly due to saturation of the spike rate). A qualitatively similar pattern has been reported previously (Rees and Palmer 1989). The neurons in C and D were “transient responders,” with a poor or absent response to long-duration unmodulated tones. Neurons of this kind maintain their tuning over a broad range of SPLs. The spike rate at a given modulation frequency varied with SPL in either a monotonic or a nonmonotonic fashion, in a manner that

could depend on modulation frequency (A and D). The BMF showed shifts of variable direction and magnitude as the SPL increased. As in Fig. 4, the tMTFs show variable degrees of a “low-pass to band-pass shift,” similar to that seen in the cochlear nucleus (Frisina et al. 1990; Rhode 1994; Rhode and Greenberg 1994).

Four examples of neurons with a suppressive region in their rMTFs are shown in Fig. 6. The suppressive region either emerges at higher SPLs (Fig. 6, A–C) as in the example of Fig. 4 or is present at all tested SPLs (Fig. 6D). The tMTFs show the same low-pass to band-pass shift seen in Fig. 5. Again, responses at modulation frequencies within the suppressive region retain considerable synchrony (i.e., show high vector strengths).

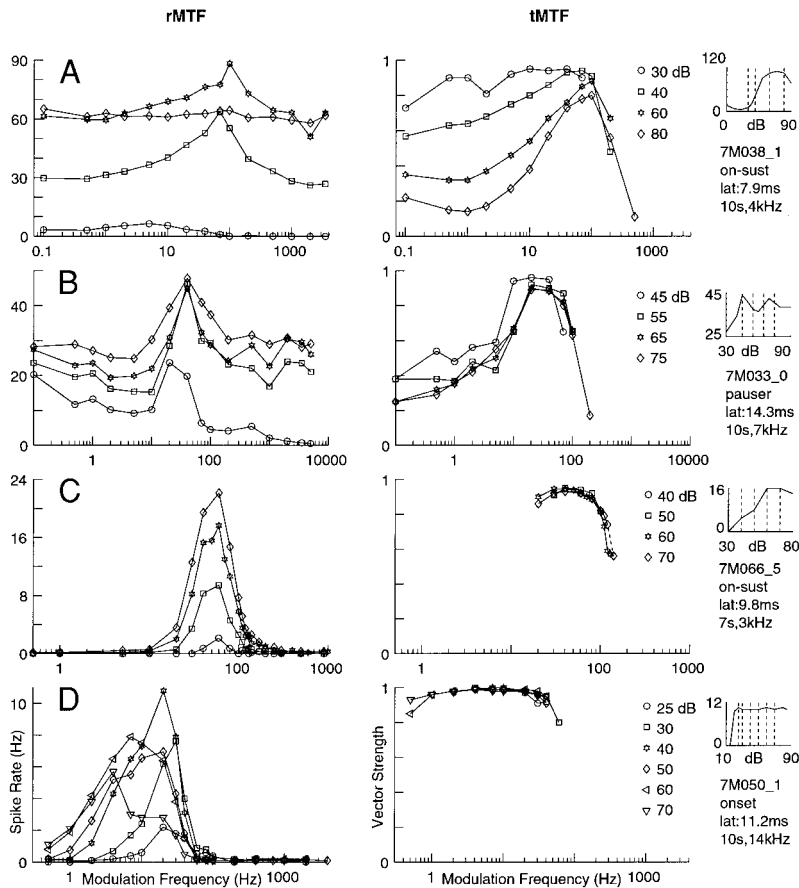


FIG. 5. Effects of varying SPL. Some rMTFs are dominated by a region of enhancement at all tested SPLs. The 1st column shows rMTFs and the 2nd column tMTFs for 4 neurons (A–D) studied with SAM tones at 100% depth, varying SPL (identified by different symbols) and fixed  $F_c$  (at or close to BF). Dashed lines in tMTFs indicate intervening responses (not plotted) with insignificant vector strength. The column at extreme right shows the RLF measured using brief tone pips, and in the text below, the neuron number, its PSTH class and minimum mean latency, and the SAM tone duration and  $F_c$  are identified. The vertical dashed lines show the different SPLs at which the SAM tone was presented. Stimulus parameters as in Fig. 2 were as follows. A: 100 ms, 10 ms, 10, 10 ms, 1. B: 200 ms, 5 ms, 10, 10 ms, 1. C: 100 ms, 10 ms, 10, 30 ms, 1. D: 100 ms, 5 ms, 20, 10 ms, 1.

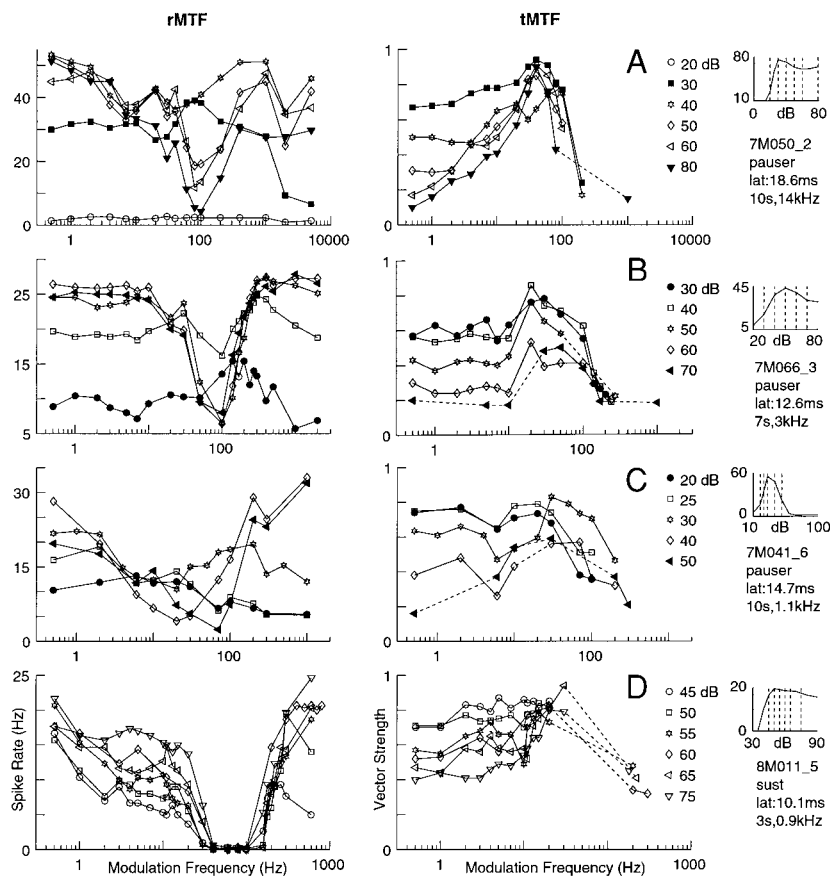


FIG. 6. Effects of varying SPL. Some rMTFs contain a region of suppression, emerging at higher SPLs or present at all tested SPLs. The format of this figure is identical to that of Fig. 5. Stimulus parameters as in Fig. 2 were as follows. A: 100 ms, 10 ms, 20, 10 ms, 1. B: 300 ms, 10 ms, 5, 50 ms, 1. C: 200 ms, 10 ms, 10, 10 ms, 1. D: 300 ms, 10 ms, 10, 10 ms, 1.

A correlation with some other neuronal response property would clearly be useful in delineating the possible mechanisms underlying the variety of rMTF changes seen with increases in SPL. No systematic relationship seems to exist between the presence of a suppressive region and the shape of the RLF (i.e., whether it is monotonic or nonmonotonic). However, it was found that the presence of a suppressive region in the rMTF at 1 or more SPLs was significantly correlated with the PSTH pattern in response to pure tones (Table 1). An rMTF region was defined as suppressive if the spike rate at modulation frequencies within that region showed a clear decrease as the modulation depth was increased (see METHODS). Neurons classified as pausers or sustained types on the basis of their PSTHs were more likely to possess rMTFs with a suppressive region, while rMTFs from neurons with onset or onset-sustained PSTHs usually only had a single region of enhancement (Fig. 5). However, it should be noted that using our aforementioned criterion, suppression can only be demonstrated clearly for

TABLE 1. rMTF type is correlated with PSTH pattern

	No Suppression	Suppression
O/Os	18	3
S/P	5*	19

Neurons that possess rate modulation transfer functions (rMTFs) with a suppressive region tend to have sustained or pauser (S/P) peristimulus time histograms (PSTHs) rather than onset or onset-sustained (O/Os) PSTHs (see METHODS).  $\chi^2$  (with a Yates continuity correction) = 16.36 ( $P < 0.001$ ). Data are shown for all 45 neurons with MTFs recorded over at least a 20-dB sound pressure level range and PSTHs recorded over a range of levels. \* All 5 neurons without suppression had pauser PSTHs.

neurons that respond well to unmodulated tones; for example, the rMTFs in Fig. 2B probably possess a suppressive region of very small magnitude between 20 and 100 Hz. Eleven of the 18 onset or onset-sustained neurons, and 2 of the 5 pauser neurons without a suppressive region in their rMTFs did not respond well to long-duration unmodulated tones. It therefore remains possible that a putative suppressive mechanism (possibly inhibition: see DISCUSSION) sharpens the high-frequency slope in rMTFs like those in Fig. 5, C and D, without actually resulting in a visible suppressive region. This is consistent with the fact that spike rates are often nonmonotonic with SPL over broad ranges of modulation frequency, even in rMTFs without suppression (e.g., Fig. 5D).

The final measure used to characterize MTFs was the mean phase of the response. Figure 7 shows the pMTFs for three neurons, whose rMTFs and tMTFs have been shown in earlier figures (Figs. 5, A and D, and 6D). A systematic increase in the phase lead (phase-advance) with increasing SPL, as illustrated in Fig. 7, A and B, was observed in the pMTFs from 30 of 56 neurons. Other neurons showed both increasing and decreasing leads over different modulation frequency ranges in their pMTFs. A decreasing phase lead (phase-delay) was sometimes seen in the pMTF in frequency regions corresponding to the decreasing high-frequency slope of the associated rMTF (especially in neurons with sustained PSTHs). This behavior was usually less systematic, but a fairly clear example is shown in Fig. 7C. The phase-advance is not always accompanied by an increase in spike rate; for example, the neuron in Fig. 7A shows a phase-advance even though its rMTFs are nonmonotonic with SPL (Fig. 5D). The phase-advance in Fig. 7B seems to be

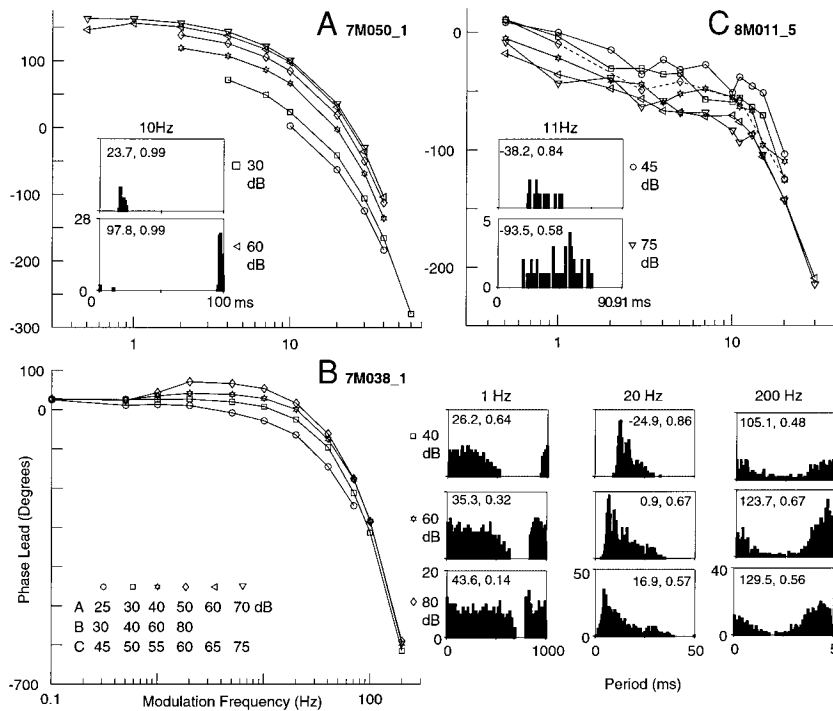


FIG. 7. Phase MTFs (pMTFs) of inferior colliculus (IC) neurons show systematic increases in lead (A and B) and lag (C) with increasing SPL. Phase-lead of the response is plotted vs. modulation frequency at multiple SPLs for 3 neurons (A–C). SPLs are identified by different symbols (symbol key in Fig. 7B). Dashed lines between points in pMTFs indicate intervening insignificant responses (not plotted). Also shown are modulation period histograms at selected modulation frequencies and SPLs for each of the 3 neurons. The text inside each period histogram gives the corresponding phase lead and vector strength. Histograms contain 100 (A and B) or 91 bins (C). See Figs. 5, A and D, and 6D for the rMTFs and tMTFs corresponding to these pMTFs. Note that phase values at very low modulation frequencies (e.g., 0.1 Hz in Fig. 7B) where only a few stimulus cycles are present, are biased (toward positive phase values) as a result of adaptation at stimulus onset.

at least partly due to the marked adaptation seen in the response that shifts the peak of the response toward the beginning of the cycle (see the period histograms). The phase advance in Fig. 7A could be due to a phasic response that occurs earlier in the cycle as the SPL increases; this could possibly also be viewed as a very rapid adaptation. There was a consistent tendency for the phase-advance to be maximum at intermediate modulation frequencies, as exemplified by the data in Fig. 7B. This implies that it is inaccurate to view the phase-advance as simply reflecting a decreased time delay at higher SPLs that mirrors the known decrease in latency with SPL for pure tones; because that explanation predicts, at least in its simplest form, a linear relationship between phase-advance and modulation frequency (i.e., the phase-advance ought then to be maximum at the highest modulation frequency). On the contrary, the phase-advance is often minimal at the highest modulation frequencies. Finally, the modulation frequency at which the maximum phase-advance in a given neuron was found was not necessarily identical to that at which its rMTF peaked.

The properties of the only offset neuron observed in this

study are shown in Fig. 8. The rMTF showed a band-pass shape, while the vector strength remained high throughout the range of modulation frequencies that elicited a response. The pMTF showed a systematic phase-delay with SPL at low modulation frequencies. The period histogram suggests that this is the result of a phasic response occurring later in the cycle as SPL increases. Consistent with its “offset” nature, the neuron fires during the falling phase of the amplitude envelope (as indicated by the negative value of the low modulation frequency asymptote). These properties are similar to the recently described properties of offset neurons from periolivary regions (Kuwada and Batra 1999).

#### MTF characteristics across the population

Of the 106 neurons responsive to SAM tones, 96 were studied by stimulating at 100% depth and with a  $F_c$  that was at or close to the BF of the neuron at all the SPLs studied. Figure 9 is a representation of the BMFs at 100% depth and different SPLs for all 96 neurons. The mean BMF (averaged across the

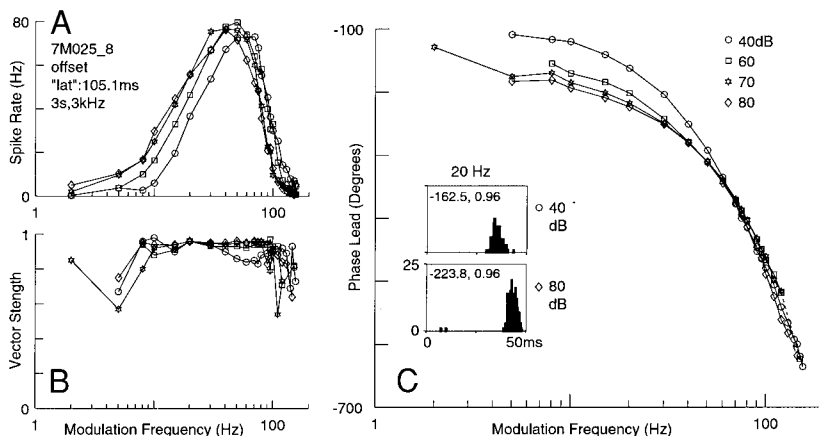


FIG. 8. One offset neuron encountered showed a peaked rMTF, excellent phase locking over a range of modulation frequencies, and a tendency to fire later during the falling phase of the amplitude envelope as the SPL increased. The rMTF (A), tMTF (B), and pMTF (C) are shown. The format for the phase plot is similar to that in Fig. 7. Cell number, PSTH class, mean minimum latency, SAM tone duration and  $F_c$  are shown in A. Each data point resulted from a single presentation of the SAM tone (10-ms rise time).

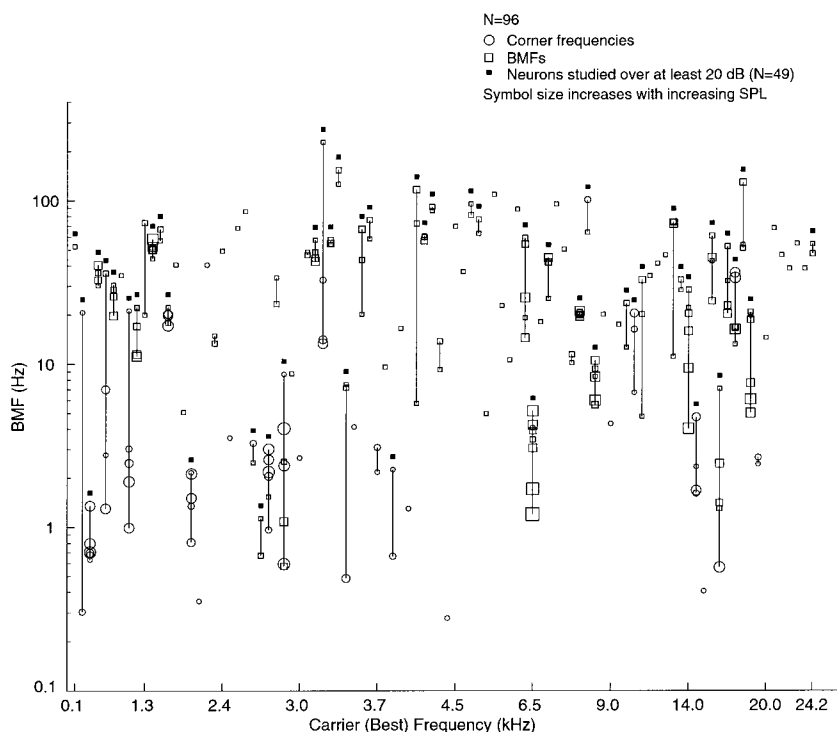


FIG. 9. Population characteristics of the BMF. BMFs from rMTFs at 100% depth and  $F_c$  at or close to BF are plotted for 96 neurons. Each vertical line joins BMFs recorded at multiple SPLs from an individual neuron; symbol size increases with increasing SPL. Neurons are arranged on the abscissa in order of increasing  $F_c$ . For rMTFs that did not show a clear primary peak, corner frequency (see METHODS) was plotted ( $\circ$ ). Neurons studied over at least a 20-dB range of SPLs are marked with a black square above the line connecting their BMFs.

different SPLs tested and including corner frequencies; see METHODS) did not show a significant correlation (Kendall's  $\tau = 0.0708$ ;  $P = 0.1533$ ) with the BF. Some interesting properties are evident in the cumulative probability distribution of the mean BMF, shown in Fig. 10A. The maximum mean BMF encountered was 140 Hz, and about 50% of the mean BMFs lay below 25 Hz. The range of variation of the BMF for an individual neuron (both absolute and relative to the mean) observed at different SPLs is also shown on the same graph. Fifty percent of the 49 neurons tested over at least a 20-dB range of SPLs showed a BMF variation larger than 66% of their mean BMF; in absolute terms, 50% of the neurons showed a range larger than 10.9 Hz. No systematic pattern was found in the variation of BMFs with SPL for individual neurons (Fig. 10B).

At least one rMTF with a suppressive region was observed in 43 of the 96 neurons. The mean WMF (averaged across SPL; see METHODS) lay between 0 and 200 Hz in most cases, with a mode near 100 Hz (Fig. 10C). No systematic pattern was found in the variation of WMFs with SPL for individual neurons (data not shown).

The minimum mean latency in response to pure tones at BF is significantly correlated (Kendall's  $\tau = -0.3276$ ,  $P = 0.0001$ ) with the mean BMF (averaged across SPL) of the neuron in response to SAM tones at 100% depth and with a carrier at or close to BF (Fig. 11). Neither measure showed a significant correlation with the BF [ $\tau$  (BF-latency) =  $-0.0288$  ( $P = 0.366$ ) and  $\tau$  (BF-mean BMF) =  $0.0329$  ( $P = 0.3479$ )]. Accordingly, a three-way test resulted in a  $\tau$  of  $-0.3269$ , only slightly different from that for the BMF-latency relationship, suggesting that the BF of the neuron was not a confounding variable.

Figure 12A reveals that cutoff frequencies (see METHODS) are markedly lower in rMTFs that possess a suppressive region, when compared with those that do not. This could be inter-

preted as an effect of the suppressive mechanism sharpening the high-frequency slope of the rMTF, thus leading to a lower rMTF cutoff frequency.

The tMTF cutoff frequency (defined here as the maximum modulation frequency at which neurons retain synchrony to the modulation frequency) can be regarded as an index of the low-pass filtering and internal noise in the system. The measure shows no significant correlation with the  $F_c$  used (Kendall's  $\tau = 0.003$ ,  $P = 0.4829$ ). Over 85% of neurons lack significant synchrony above 300 Hz (Fig. 12B). Cutoff frequencies in the IC are thus substantially lower than those found at lower levels like the cochlear nucleus and the lateral superior olive (Joris and Yin 1998; Rhode and Greenberg 1994; both in the cat).

Some other aspects of the vector strength transformation between the cochlear nucleus and the IC are shown in Fig. 13. As shown in Fig. 13A, the maximum synchrony found in responses from almost all IC neurons ( $0.907 \pm 0.009$ , mean  $\pm$  SE,  $n = 96$ ) is clearly more than the means reported for any neuron type in the cochlear nucleus, with the possible exception of  $O_1$  units (Rhode 1994; Rhode and Greenberg 1994). Also, responses from neurons that possess BMFs (i.e., the rMTF shows a clear peak, see METHODS) at all SPLs studied are almost maximally synchronized at the BMF; i.e., the mean difference between the vector strength at BMF and the maximum vector strength is close to zero (e.g., MTFs in Figs. 2, A-C, and 5). In contrast, neurons that possess corner frequencies show larger mean differences between the vector strength at their rMTF peak (BMF or corner) and the maximal vector strength in the associated tMTF (for example, see MTFs in Figs. 3, 4, and 6).

The examples in Figs. 2 and 3 suggest that IC neurons show considerable phase locking (i.e., high vector strength) at relatively low depths of modulation. This is confirmed in Fig. 13B, where the lowest modulation depth at which a significant

vector strength was found is plotted against the value of the vector strength for the 31 neurons whose MTFs were recorded at multiple depths. Comparing this to modulation depth-vector strength functions recorded from the cat cochlear nucleus population (Rhode 1994: Fig. 13) confirms that vector strengths of IC neurons at low depths are more than those for almost all

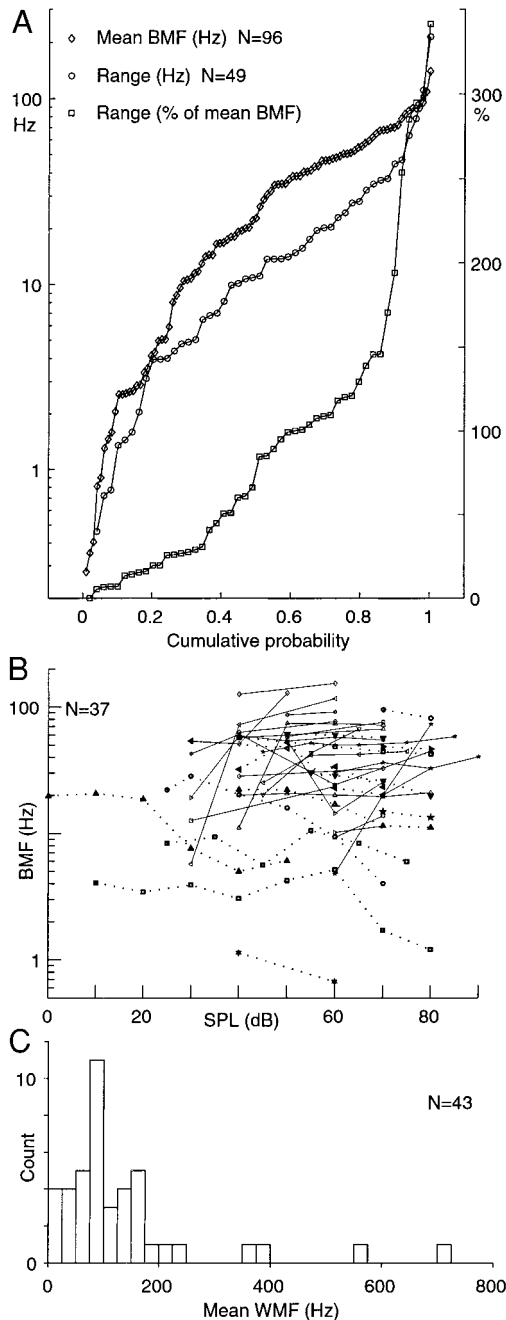


FIG. 10. Most mean BMFs and WMFs lie below 100 Hz, and over half the neurons show a BMF range larger than 66% of their mean BMF. *A*: mean BMF, absolute range (both plotted on the *left ordinate*) and range as a percentage of the mean BMF (*right ordinate*) plotted vs. their cumulative probability. *B*: BMF variation with SPL did not show any systematic pattern. Lines join BMFs recorded from a single neuron at different SPLs; data are shown for neurons that showed peaked MTFs at all (>1) tested SPLs. Dotted lines connect data from neurons whose BMF at the highest SPL is lower than that at the lowest SPL. *C*: histogram of mean WMFs from neurons with a suppressive region in their rMTFs.

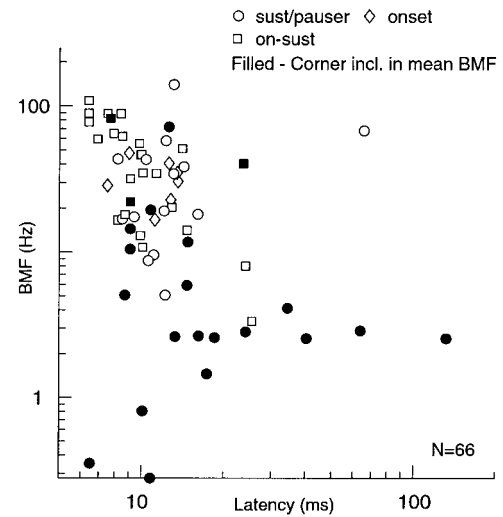


FIG. 11. Neurons with shorter minimum latencies have higher mean BMFs. Mean BMFs are plotted vs. minimum mean latency for neurons whose minimum latency was identified using a pure tone RLF. Symbols indicate PSTH type: onset (◊), sustained or pauser (○), and onset-sustained (◻). Filled symbols have corner frequencies included in the mean BMF calculation (see legend to Fig. 9).

cochlear nucleus neuron types (onset-choppers being the exception).

One pMTF each from 95 of the 96 neurons is plotted in Fig. 14A to display the population characteristics of the pMTFs. The positive low modulation frequency asymptotes indicate a tendency for most neurons to fire on the rising portion of the sinusoidal envelope. A straight line was often a good fit to the high-frequency (greater than or equal to 100 Hz; in our observations, responses in this range usually showed small or absent SPL-dependent phase shifts) phase responses; the maximum and minimum values of the slope (see legend) across all neurons were 15.6 and 6.09 ms. These may be interpreted as time delays (e.g., Anderson et al. 1971; but also see Ruggero 1980) that contain possible contributions from the fixed delays and filtering properties of the system.

Thirty of the 56 neurons studied at multiple SPLs showed a systematic phase-advance as the SPL increased. This is shown in Fig. 14B, where phase leads (at the modulation frequency at which the largest phase-advance was found) are plotted as a function of SPL for each of the 30 neurons. The distribution of the maximum phase-advance (i.e., the range of each of the 30 functions in Fig. 14B) per 10-dB rise in SPL is shown as a histogram in Fig. 14C. The mean maximum increase was  $23.81^\circ$  per 10 dB (SE, 1.78). This may be larger than similar increases reported from responses at lower levels in the auditory pathway (see DISCUSSION).

#### Varying carrier frequency

We also investigated in 34 neurons the effects of stimulating different parts of the frequency response area by varying the  $F_c$  of the SAM tone. Illustrative examples are shown in Figs. 15 and 16. The examples in Fig. 15 show that band-pass rMTFs derived from an individual neuron can have different shapes; i.e., they are not scaled versions of each other. The BMFs can show shifts of varying magnitude and direction. The spike rate varies with carrier frequency in a manner that is roughly

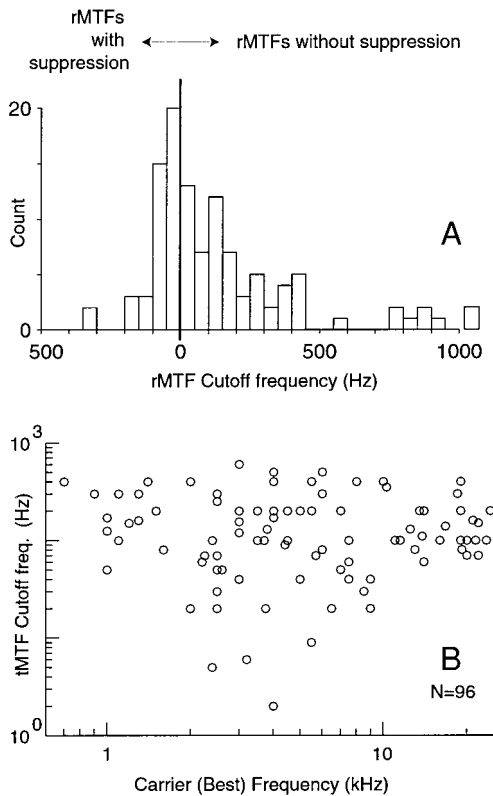


FIG. 12. Properties of the cutoffs of rMTFs and tMTFs. *A*: histograms of mean cutoff frequencies of rMTFs with ( $n = 43$ ) and without ( $n = 67$ ) suppression (averaged across SPL, from individual neurons) plotted, respectively, to the left and right of the dark line at 0 Hz; 14 neurons contributed to both sides, and 2 points (1,970.43 and 1,269.4 Hz) are plotted outside the range for clarity. It is unlikely ( $P < 0.001$ , Mann-Whitney  $U = 5.14$ ) that the 2 histograms arose via random sampling from a common distribution. *B*: tMTF cutoff frequency (maximum modulation frequency at which a significant vector strength was found) vs.  $F_c$  (at or close to BF). One data point (125 Hz  $F_c$  and 120 Hz cutoff frequency) was omitted for clarity; 5,149 points were tested for significance while creating this plot.

consistent with the frequency response area. Finally, as discussed earlier in the context of Fig. 4, one might imagine that suppressive regions could result from the positioning of the AM sidebands over inhibitory regions of the frequency response area. This explanation might also predict that as  $F_c$  is varied, suppressive regions would show shifts of the same magnitude as the shift in  $F_c$ . This does not seem to be the case (Fig. 16, A–C).

Figure 16*D* shows data from a neuron whose rMTF seems to lack suppressive regions when stimulated well away from its BF. Because suppressive regions can emerge at higher SPLs (see Fig. 6*A* for MTFs at multiple SPLs from the same neuron), this raises the issue of whether the effects of  $F_c$  variation can be explained in part on the basis of the SPL relative to threshold at each  $F_c$ . We do not yet have sufficient data to address this question.

The tMTF variation with  $F_c$  is roughly similar to the variation observed in the same neuron when SPL is varied. Thus when  $F_c$  is varied at a particular SPL, tMTFs appear to vary over the same range of values seen when stimulated between threshold and that particular SPL at BF. For example, the neuron in Fig. 15*C* showed almost invariant tMTFs as the SPL was varied (data not shown, but similar in this respect to Fig.

5, *C* and *D*); this was mirrored in the tMTF invariance with  $F_c$ . Similarly, the neuron in Fig. 16*A* showed a large drop in vector strength at low modulation frequencies as SPL increased (Fig. 4); the tMTF in Fig. 16*A* shows the same property as  $F_c$  became closer to BF. However, the details of the variation within this range did not show any clear pattern, especially at higher levels; for example, the vector strength did not show a consistent relationship either with  $F_c$  or with spike rate. These results are similar to those reported from the cochlear nucleus (Rhode 1994). Finally, no systematic changes in the pMTFs were found when  $F_c$  was varied.

One scheme for the generation of spike rate tuning to modulation frequency in the midbrain (Langner 1981) for low carrier frequencies predicts a linear relation between

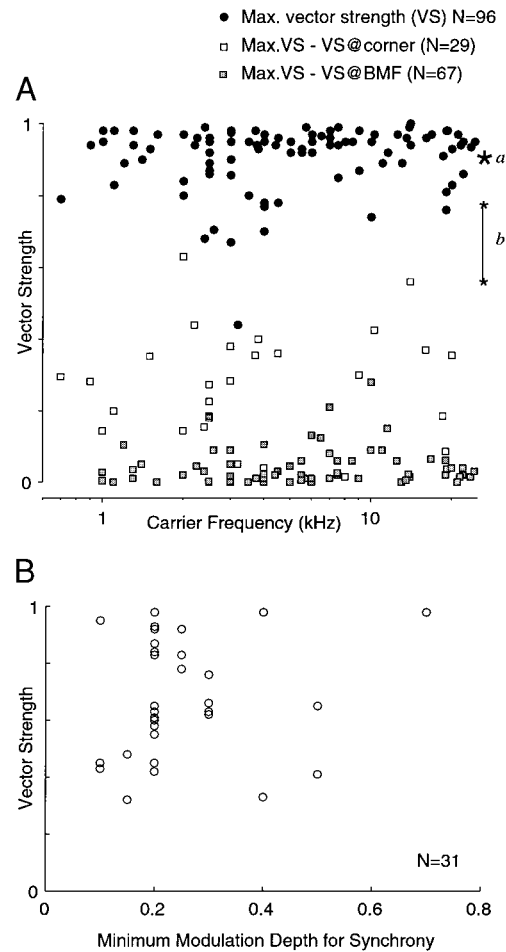


FIG. 13. IC responses are more synchronized than cochlear nucleus responses. *A*: the maximum vector strength (●), and the mean difference (across different SPLs, at 100% depth) between the vector strength at the rate BMF and the maximum vector strength (□ and ▣) are plotted vs.  $F_c$  (at or close to BF). □, the 29 neurons that have at least 1 rMTF with a corner frequency (instead of a BMF); ▣, the other 67 neurons. One □ (BF of 125 Hz, maximum vector strength of 0.85 and a difference of 0.04) was omitted for clarity. The mean maximum vector strengths are indicated by asterisk *a* for the IC population (0.907,  $n = 96$ ) and by range *b* (0.56–0.78) for the various neuron types in the cochlear nucleus (Rhode and Greenberg 1994, cat). *B*: minimum depth of stimulation (across all modulation frequencies tested) that elicited a significantly synchronized response plotted vs. the vector strength of the response for neurons studied by varying depth systematically while keeping  $F_c$  (at BF) and SPL constant. Note that at 20% modulation depth, stimulus amplitude varies from –1.9 dB below to 1.6 dB above mean SPL; at 40%, the corresponding range is –4.4 to 2.9 dB.

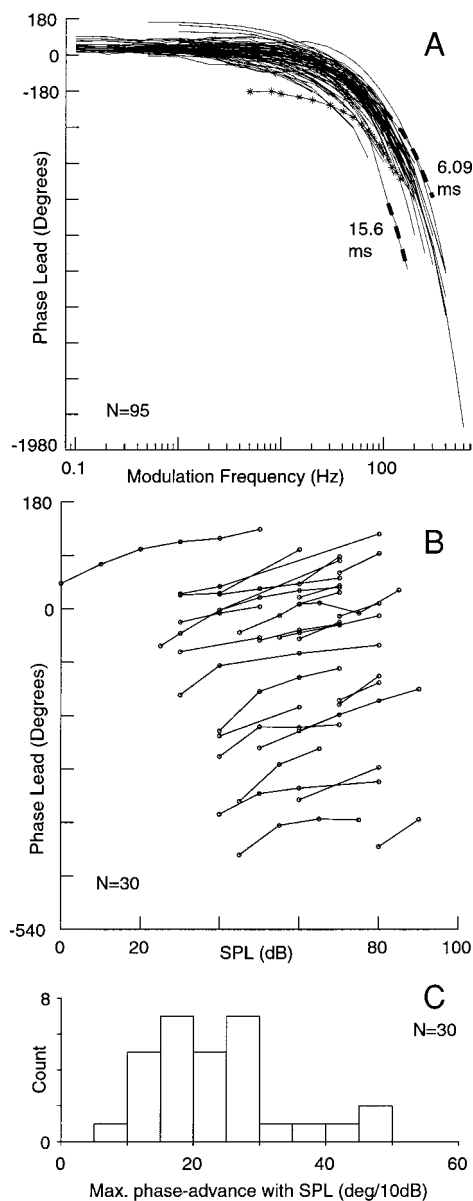


FIG. 14. Population characteristics of the phase response. *A*: the pMTF at a single SPL (usually the highest of all SPLs tested) plotted vs. modulation frequency for 95 neurons; 1 neuron that exhibited response synchrony at fewer than 4 modulation frequencies was excluded. Thick dashed lines show straight line fits to the pMTFs with maximum and minimum slope above 100 Hz (see text). Slope values (divided by  $360^\circ$  and multiplied by  $-1$ ) are shown in milliseconds. The curve with asterisks is from the offset neuron in Fig. 8. *B*: a clear phase advance with rise in SPL was seen in 30 of 56 neurons studied at multiple SPLs (with  $F_c$  at or close to BF). Each line joins phase leads measured at different SPLs (at the particular modulation frequency that produced the maximum advance) from a single neuron. *C*: histogram of the phase advance per 10 dB rise in SPL for the sample in *B*.

$1/\text{BMF}$  and  $1/F_c$  (and therefore a monotonic relationship between BMF and  $F_c$ ). Figure 17, *A–E*, show data from the five cells (in this study) with peaked rMTFs that were studied at multiple low carrier frequencies ( $<5$  kHz) within their frequency response area. No systematic pattern was found for BMF shifts with  $F_c$ . Plotting these shifts as functions of  $1/\text{BMF}$  versus  $1/F_c$  (Fig. 17*F*) confirms the lack of any systematic linearity in our data.

## DISCUSSION

The present study provides detailed descriptions of the variations in spike rate, synchrony, and phase of spike discharges of single neurons in the IC in response to variations in modulation depth, SPL, and  $F_c$  of SAM tone stimuli at a range of modulation frequencies. The systematic (and often dramatic) changes seen in all these response measures when stimulus parameters are varied have been described in RESULTS. We now discuss the implications of these findings both for the neural mechanisms generating AM responses in the IC and for performance in psychophysical tasks involving auditory temporal processing.

### *Excitation and inhibition together create the IC rMTF*

Most neuron types in the cochlear nucleus predominantly show poorly tuned or nonexistent variations in spike rate (i.e., low-pass or flat rMTFs) with modulation frequency (Frisina et al. 1990; Rhode and Greenberg 1994). Information about the modulation frequency is instead present in the response component locked to the modulation frequency (i.e., a “temporal code” rather than a “rate code”). In contrast, neurons in the IC show large variations of spike rate with modulation frequency. These variations have previously been reported to result in rMTFs of varied complex shapes, with band-pass rMTFs that possess a BMF being the most common variety (Heil et al. 1995; Langner and Schreiner 1988; Rees and Palmer 1989). The data in this paper show that rMTFs are composed of regions of enhancement and suppression, where the spike rate increases or decreases, respectively, with an increase in the modulation depth of the SAM tone stimulus. In particular, almost all IC rMTFs could be described by some combination of a primary and a secondary region of enhancement and an intervening region of suppression (Fig. 18), with these regions present to varying degrees in individual rMTFs. The regions of enhancement have band-pass shapes, with the low-frequency region of enhancement (E1) forming the primary peak (BMF or corner frequency), and the much less commonly found high-frequency region (E2) the secondary peak. The region of suppression creates the band-suppressive shape, with its trough forming the WMF.

In the spirit of most current models (see *Mechanisms generating the secondary region of enhancement*), we speculate that the primary region of enhancement is primarily the result of a transformation of excitatory inputs to the IC. This seems plausible because at least some rMTFs in the mustache bat IC remain band-pass after iontophoretic application of various inhibitory blockers (Burger and Pollak 1998). A similar result has been reported for onset neurons in the mustache bat dorsal nucleus of the lateral lemniscus (DNLL) (Yang and Pollak 1997), supporting the notion that purely excitatory mechanisms are capable of creating band-pass rMTFs. However, it is possible that both in the IC and the DNLL, the inputs themselves show band-pass spike rate tuning (see next section); this could also account for the minimal effect of inhibitory blockers on the band-pass tuning of IC rMTFs. On the other hand, it seems very likely that inhibitory inputs shape the rMTF by creating the region of suppression. In addition, inhibition may sharpen the high-frequency rolloff of the primary region of enhancement without resulting in a visible region of suppression (as a result of the low firing rate in response to unmodulated tones:

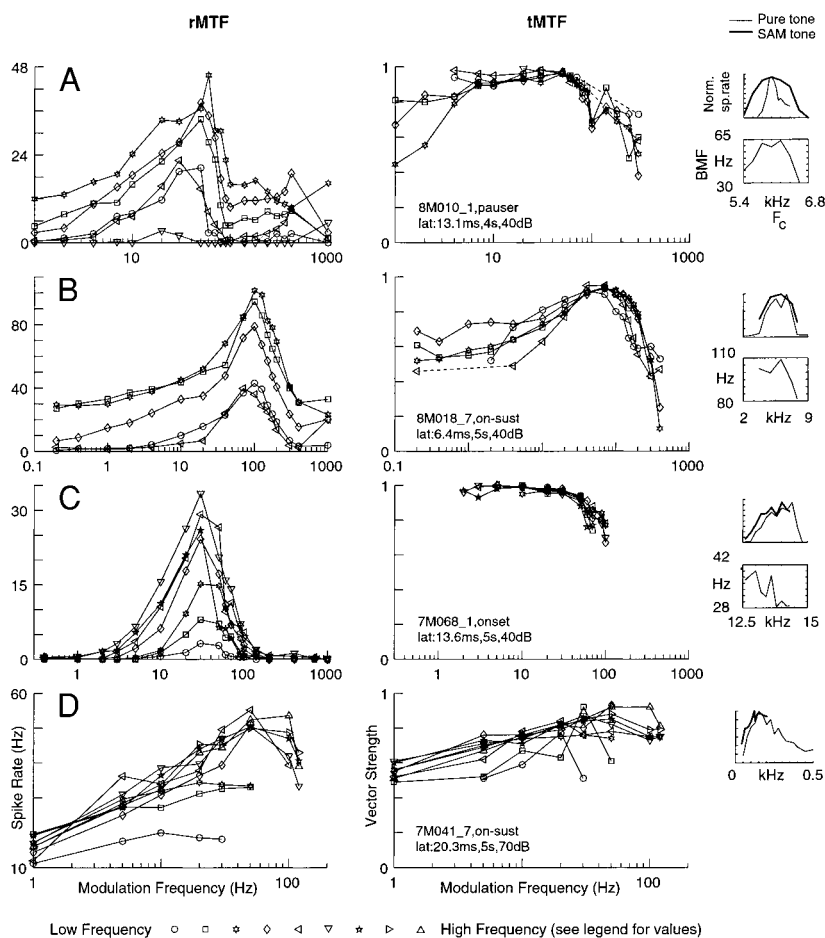


FIG. 15. Illustrative examples of neurons with band-pass rMTFs studied at multiple  $F_c$ s. rMTFs (1st column) and tMTFs (2nd column) shown for 4 neurons (A–D) studied with SAM tones at 100% depth, constant SPL and varying  $F_c$  (identified by different symbols). Dashed lines indicate intervening responses (not plotted) with insignificant vector strength. The top plot at extreme right in each row shows the spike rate vs. frequency function measured using brief tone pips (thin line) and SAM tones (thick line, modulation frequencies of 60, 100, 30, and 30 Hz in A, B, C, and D respectively), both normalized to their respective peak values and measured at SPLs within 10 dB of each other. The bottom plot shows the dependence of BMF on  $F_c$ . The rMTFs in D did not have a clear BMF at many  $F_c$ s, and the BMF- $F_c$  plot was omitted. The cell number, PSTH class, mean minimum latency, SAM tone duration and SPL are given inside the tMTF axes. Carrier frequencies shown in A: 5.4, 5.6, 6.0, 6.4, 6.6, and 6.8 kHz. B: 3.5, 4.5, 5.5, 6.5, and 7.0 kHz. C: 12.6, 12.8, 13.0, 13.4, 13.6, 14.0, and 14.3 kHz. D: 40–200 Hz at intervals of 20 Hz. rMTFs and tMTFs at a few intervening  $F_c$ s followed the trend shown by those at flanking  $F_c$ s and were omitted for clarity; however, points from these omitted MTFs are included in the panels at extreme right. SAM tone rise time and number of presentations in A: 30 ms, 1; B: 30 ms, 2; C: 30 ms, 1; D: 10 ms, 1.

see discussion of Table 1 in RESULTS). The properties of the secondary region of enhancement remain unclear; potential mechanisms that could generate it are discussed in a later section.

In the next few sections, various issues related to the observed MTFs are discussed in greater detail. However, it must be emphasized that in the absence of much pertinent data about intrinsic properties and input patterns of IC neurons, as well as the AM response properties of inputs to the IC, much of this discussion must remain speculative.

#### Is rate tuning created *de novo* in the IC?

With contralateral sound presentation, the excitatory afferents to the IC that are likely to be active include those from the contralateral cochlear nucleus (in particular, stellate cells in the ventral cochlear nucleus, and fusiform cells in the dorsal cochlear nucleus), contralateral lateral superior olive (LSO), ipsilateral medial superior olive (MSO), and possibly, the contralateral IC (Moore et al. 1998; Oliver and Huerta 1991). Most neurons in the cochlear nucleus do not show much rate tuning. Although little is known about the response properties of MSO neurons to SAM tones, it appears from preliminary data that neurons in the LSO possess, in addition to tuned tMTFs, rMTFs that show systematic changes of spike rate with modulation frequency (Thornton SK and Semple MN, unpublished observations). LSO rMTFs can show peaks (BMFs); the majority of BMFs seem to lie below 400 Hz. Thus rMTFs of IC

neurons may at least in part reflect the rate tuning present in their rate-tuned inputs (e.g., from the LSO); low-pass filtering of the inputs (which are also phase locked to the modulation frequency) may potentially account for the lower BMFs in the IC. However, the extent to which different inputs overlap in their projections onto single IC cells is not clear. It therefore remains plausible that at least some of the rate tuning seen in the IC emerges as a result of collicular processing.

#### Coincidence detection mechanisms may create the primary region of enhancement

Mechanisms clearly exist in the auditory midbrain to create tuned rMTFs from inputs (e.g., from the cochlear nucleus) that show varying amounts of synchrony to the modulation frequency, but little or no spike rate changes with modulation frequency. One candidate scheme suggests that band-pass rMTFs (regions of enhancement) seen in IC neurons result from coincidence detection of synchronized excitatory inputs (Hewitt and Meddis 1994). In this model, the IC neuron is considered to be a coincidence detector that fires maximally when its inputs (from the cochlear nucleus) are maximally synchronized, thus converting the peak in the input tMTFs to a rMTF peak in the IC. Such a model can reproduce some aspects of the previously reported data, including the flattening of some rMTFs at high SPLs (Rees and Palmer 1989) (Fig. 5A in present study: because the coincidences generated by the high input spike rates at high SPLs fire the coincidence detector

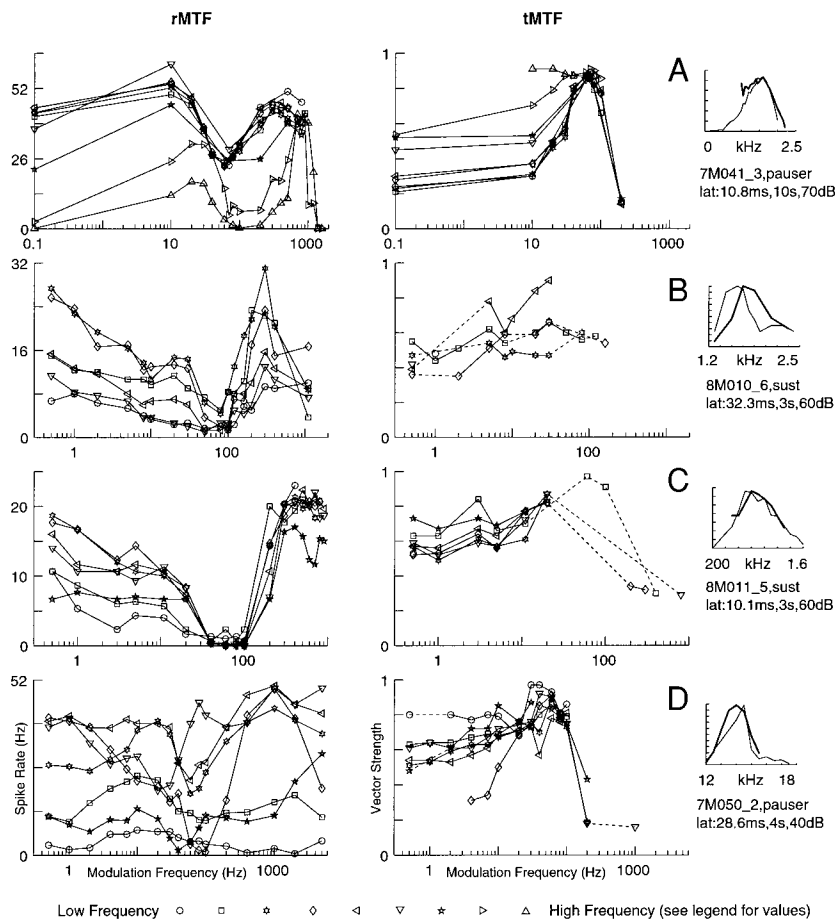


FIG. 16. Illustrative examples of rMTFs containing a suppressive region, from 4 neurons studied at multiple  $F_c$ s. The format is similar to that of Fig. 15, except for the absence of the BMF- $F_c$  plots. rMTFs and tMTFs at a few intervening  $F_c$ s were omitted for clarity.  $F_c$ s presented in A: 1, 1.1, 1.25, 1.35, 1.45, 1.6, 1.8, 2, and 2.15 kHz; B: 1.3, 1.6, 1.8, 2, 2.3, and 2.6 kHz; C: 0.5, 0.6, 0.8, 0.9, 1, 1.1, and 1.3 kHz; D: 12, 12.5, 13, 13.5, 14.5, and 15.5 kHz. SAM tone rise time and number of presentations in A: 10 ms, 1; B: 50 ms, 1; C: 10 ms, 1; D: 10 ms, 1. Modulation frequencies used for  $F_c$ -spike rate function at extreme right were 10, 0.5, 0.5, and 1 Hz in A, B, C, and D, respectively.

independent of the synchrony in the inputs). However, in comparison to the tMTF peaks in the inputs from the cochlear nucleus (80–520 Hz: Frisina et al. 1990, gerbil; mean = 330

Hz: Rhode and Greenberg 1994, cat) and the lateral superior olive (200–600 Hz: Joris and Yin 1998, cat), IC BMFs appear to span a much lower range (0–100 Hz: Fig. 10A). In other

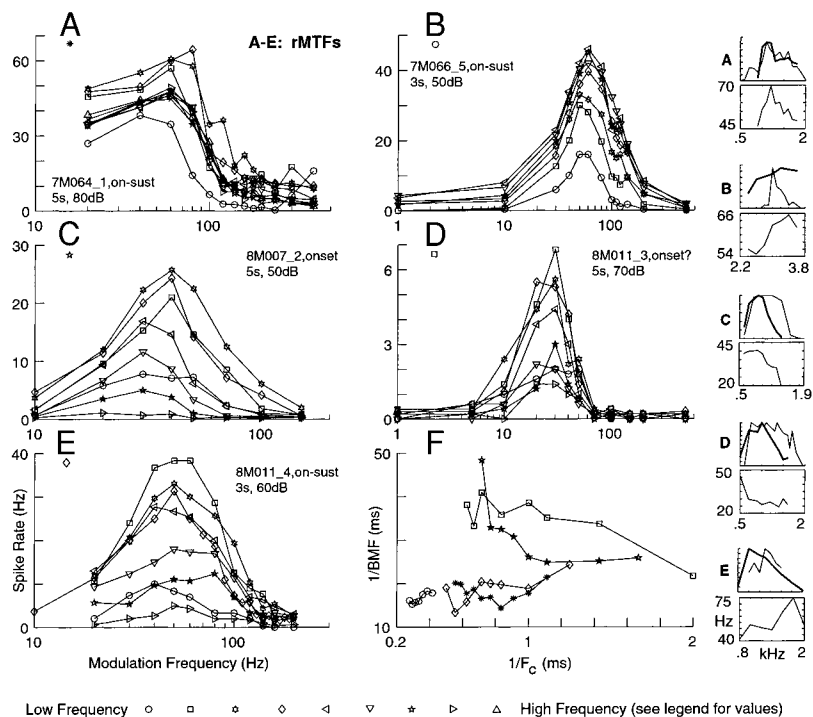


FIG. 17. BMFs from rMTFs at different  $F_c$ s within the frequency response area do not show any consistent pattern. A–E: rMTFs from the 5 low-frequency neurons in the population with clearly defined rMTF peaks are displayed. The format is similar to that in Fig. 15, except for the absence of the tMTFs. The PSTH pattern in D was measured at a single SPL. F:  $1/\text{BMF}$  vs.  $1/F_c$  plots are not linear. Each of the 5 lines joins BMFs at different  $F_c$ s for a single neuron (extracted from the rMTFs in A–E and identified by the symbols next to the letters A to E).  $F_c$ s presented in A: 0.9–1.8 kHz at intervals of 0.1 kHz; B: 2.4–3.6 kHz at intervals of 0.2 kHz; C: 0.6–1.4 kHz at intervals of 0.1 kHz, 0.8 kHz was omitted; D: 0.5, 0.7, 0.9, 1, 1.2, 1.4, 1.5, 1.6 kHz; E: 0.8–1.4 kHz at intervals of 0.1 kHz, 1.6, 1.8, and 2 kHz. SAM tone rise time and number of presentations in A: 10 ms, 1; B: 50 ms, 1; C: 30 ms, 2; D: 50 ms, 1. E: 50 ms, 1. Modulation frequencies used for  $F_c$ -spike rate function in extreme right were 60, 50, 40, 20, and 60 Hz in A, B, C, D, and E, respectively.

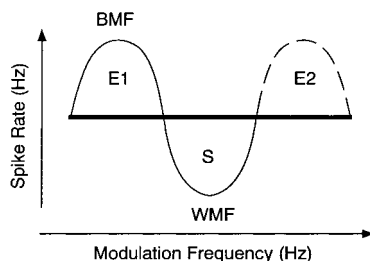


FIG. 18. IC rMTFs can be described by some combination of a primary (E1) and a secondary (E2) region of enhancement and an intervening region of suppression, with these regions present to varying degrees in individual rMTFs. E1 forms the primary peak (BMF or corner frequency) and E2 the secondary peak. Modulation frequency in S where the spike rate reaches a minimum is the WMF. Horizontal line shows response to an unmodulated tone. E2 is outlined using dashed lines to indicate that it was only occasionally found, and that its properties remain unclear (see text).

words, rMTF peaks in the IC do not seem to be equal to tMTF peaks in their inputs, as specified in the model. Other mechanisms (possibly including some combination of a stage of low-pass filtering, inhibitory inputs causing a sharper high-frequency rolloff, or intrinsic cellular properties) may need to be included to account for the data. However, it seems *prima facie* possible that such a model (with the addition of inhibitory inputs to create the region of suppression: see next section) might serve as a good first attempt to reproduce various other aspects of AM responses reported in this paper. An extensive modeling study would also offer insight into input patterns and cellular properties that could generate the diversity of MTF characteristics seen in the IC.

An alternate scheme that has been proposed to explain peaked rMTFs also treats the IC neuron as a coincidence detector. However, the structure of this model is very different from that above: it is posited that a cross-correlation analysis is performed by neurons that detect coincidences between spike trains synchronized to the modulation frequency and carrier frequency, respectively, and delayed by different small time periods (Langner 1981). For various reasons, such a scheme seems less attractive than the model discussed above. First, the model only works at low carrier frequencies (<5 kHz), where IC inputs retain synchrony to the carrier; thus requiring that a different mechanism be invoked to explain the rMTFs in neurons at high carrier frequencies (which seem, at least qualitatively, similar to those at low carrier frequencies). Second, there is little evidence that inputs of the kind required by the model actually exist. Finally, the evidence supporting one of the key predictions of the model (that there is a linear relationship between  $1/\text{BMF}$  and  $1/F_c$ ) is unclear; a linear relationship with positive slope has been reported from one multiple unit in the cat IC (Langner and Schreiner 1988), while a negative slope has been reported to exist for about 10% of neurons in the guinea fowl midbrain nucleus (Langner 1983). No systematic linear relationship was found for the five neurons examined using low carrier frequencies in the present study.

#### *Inhibition (tonic or phasic) creates the region of suppression*

The IC receives multiple inhibitory inputs that are expected to be driven by monaural input to the contralateral ear; some of the prominent ones include the DNLL bilaterally, the ipsilateral ventral nucleus of the lateral lemniscus, the opposite IC,

and inhibitory interneurons within the IC (Moore et al. 1998; Oliver and Huerta 1991). It therefore seems plausible to assume (in the absence of more direct evidence) that the suppressive region is a result of inhibition. The magnitude of the suppressive effect is dependent on the modulation frequency, with maximum suppression occurring at the WMF. In one simple scheme, this dependence on modulation frequency may reflect the fact that the net inhibitory input may show rate tuning and thus possess a BMF. Assuming that the inhibitory effect is simply proportional to the mean spike rate of the inhibitory inputs (i.e., the inhibitory inputs act in a tonic manner), the maximal inhibitory effect would then be exerted at the BMF for the inhibitory inputs. The WMF in the rMTF of an IC neuron would thus correspond to the BMF of its inhibitory inputs. Preliminary results (Thornton SK and Semple MN, unpublished observations) suggest that neurons in the gerbil DNLL can indeed show peaked rMTFs with peaks in the same frequency range as the IC WMFs (0–200 Hz; Fig. 10C). The inputs from the opposite IC and the intra-IC inhibitory input will have lower BMFs; these could result in the WMFs in the lower end of the range.

In an alternative scheme, even if the mean rate of the inhibitory inputs is independent of modulation frequency, time delays between excitatory and inhibitory inputs (both of which are synchronized to the modulation frequency) can lead to suppression whose magnitude depends on modulation frequency. For example, if the inhibitory input is slightly delayed (by a constant amount, independent of modulation frequency) with respect to excitation, then when the modulation period is equal to the delay period, the inhibitory input from a preceding cycle will overlap maximally with the excitatory input from a succeeding cycle, thus leading to a minimum response from the output cell (see Grothe 1994 for such a model proposed for the MSO of the mustached bat). At higher frequencies, the spike rate may recover back to a higher value because the inhibitory input could then become unsynchronized to the modulation frequency. The precise rMTF shapes predicted by such a scheme are likely to depend on a quantitative specification of the model.

Most of the inhibitory inputs to the IC remain synchronized in the frequency range where IC WMFs lie (0–200 Hz). However, the combination of the multiple inputs that sum at potentially random phase (with respect to each other), the longer time scales associated with inhibitory postsynaptic potentials, and pre- and postsynaptic filtering could result in a tonic effect. It is thus not possible to distinguish between the two (and other more complex) postulated schemes for the mechanisms of inhibitory effect at this point.

Since there is no direct evidence implicating inhibitory input as the suppressive mechanism, it seems worthwhile to consider mechanisms that could potentially explain the drop in response below that to the unmodulated tone without invoking inhibition. When the input spike rate is high, and a coincidence detector has a short refractory period, allowing it to fire multiple times during a single cycle, synchronized inputs may actually lead to a lower spike rate than inputs that occupy more of the cycle period, thus allowing the detector to fire more than once in each cycle (see Reed and Durbeck 1995). However, this mechanism seems inconsistent with the almost complete suppression of the response in many cases (Figs. 3 and 6). Further, it is not clear whether such a scheme (or for that

matter, other schemes based purely on excitatory inputs) could account for the fact that the magnitude of suppression depends on modulation frequency.

#### *Mechanisms generating the secondary region of enhancement*

The secondary region of enhancement is of some interest, if only because it generates a secondary peak at a high modulation frequency in the rMTFs of eight neurons in our population. In six of these neurons, this was demonstrably not a result of the sidebands moving into local peaks in the frequency response area, as it is seen in response to stimulation with a wide range of carrier frequencies (and SPLs). Since responses at the secondary peak were not synchronized to the modulation frequency, it seems unlikely that it could be created by coincidence detection mechanisms. There are many other ways in which such peaks could be generated. A tuned inhibitory input may create a suppressive region in the middle of an otherwise low-pass or band-pass rMTF, which would create a secondary peak. Alternatively, these could reflect high-frequency peaks in the input rMTFs. Yet another possibility is that they are a result of spectral interactions above the cochlear nucleus. For example, if the dynamic range of individual inputs to the IC is narrow, there will be a stronger net input at higher modulation frequencies when the sidebands excite different inputs as a result of their becoming more spectrally separated. If this were the case, these secondary peaks would be invariant to changes in relative phase between the three components of the SAM stimulus. Finally, at high modulation frequencies (where the lower and upper sideband are well separated from each other), it is possible that the rebounds in response actually reflect differences in the speaker transfer function at the carrier and the two sideband frequencies (as noted in METHODS, the compensation was performed prior to modulation with the  $F_c$  as reference). MTF shapes resulting from this are likely to change as the  $F_c$  is varied. Differentiating between these various possibilities awaits further experimentation.

#### *tMTFs*

The low-pass to band-pass shift with increasing SPL seen in IC tMTFs is similar (at least qualitatively) to that observed in the cochlear nucleus (Frisina et al. 1990; Rhode 1994; Rhode and Greenberg 1994). This is probably because at low modulation frequencies (where the stimulus changes minimally over a time-scale equivalent to the "integration time" of the neuron), the response in neurons that do not respond in a phasic manner can be roughly viewed as tracking the instantaneous SPL of the stimulus (e.g., Cooper et al. 1993). As the SPL increases, the stimulus remains suprathreshold for longer durations during each cycle, thereby eliciting a broader response (at least at low modulation frequencies). The period histograms in Fig. 7B suggest that this indeed seems to be the case, as has been pointed out by previous investigators (e.g., Hewitt and Meddis 1994; Rees and Palmer 1989). Neurons tend to maintain their vector strength at the peak of the tMTF at all SPLs, thus generating the band-pass shape at high SPLs. The vector strength changes appear to be independent of the associated spike rate changes in the response, even at low modulation frequencies (Figs. 5 and 6). The considerable variability in the

tMTF variations with SPL and  $F_c$  seen in the IC (present study) is probably at least in part, a reflection of the similar variety of changes seen in the different neuron types of the cochlear nucleus (which directly or indirectly, are the source of the inputs to the IC).

Compared with neurons in the cochlear nucleus, IC neurons also show enhanced synchrony (higher vector strengths: Fig. 13) and lower cutoff frequencies for synchronization (Fig. 12). The enhancement is consistent with previous results that indicate a progressive increase in the maximum vector strength values achieved at higher levels (lateral superior olive > cochlear nucleus > auditory nerve) in the auditory pathway (Joris and Yin 1998). The increased synchrony (relative to that in cochlear nucleus neurons) found in IC neurons (present study: Fig. 13) is consistent with the predictions of the hypothesis that the IC neuron acts as a coincidence detector (Burkitt and Clark 1999; see also Wang and Sachs 1995). However, other possible explanations include 1) inhibition raising the threshold for spiking (and thus decreasing the range of phases over which the neuron discharges), 2) successive adaptation making the period histogram narrower by speeding up the response decay within each cycle, and 3) a decreased spontaneous rate at higher levels of the auditory pathway (possibly as a result of anesthesia). The reasons behind the lower cutoff frequencies measured in the IC are not entirely clear; but factors similar to those indicated in a preceding section for the loss of synchrony in inhibitory inputs may play a role. However, the presence of rate tuning in the IC also means that a lack of sufficient response at high modulation frequencies often limits the detection of any possible synchrony. This, combined with the fact that high modulation frequencies were not sampled with an aim to extract the cutoff precisely, means that the measured values are likely to be an underestimate of the true cutoff values.

#### *pMTFs*

A systematic phase-advance was observed in the pMTFs of 30 of 56 neurons as the SPL increased. Phase-advances with SPL have also been documented for the auditory nerve, spherical and globular bushy cells in the cochlear nucleus, the medial nucleus of the trapezoid body, and the lateral superior olive (all in the cat) where the mean phase advances observed are 4.07, 4.1, 6.0, 7.3, and 9.97°/10 dB, respectively (Joris and Yin 1998). All of these values are substantially lower than the mean maximum increase of 23.81°/10 dB for IC neurons in the present study. However, the values for nuclei other than the IC result from measurements made at a predetermined modulation frequency (100 Hz) that was not necessarily the frequency at which the maximum advance would be seen. Nevertheless, the seemingly systematic increase in the phase advance observed at succeeding levels in the auditory pathway may be the result of an incremental advance at each stage of processing. The phase-advance seems to be at least partly due to a marked adaptation seen in the response during each cycle. This also seems to be the basis for this phenomenon in the auditory nerve (Joris and Yin 1992; for a similar explanation for phase-advance in V1 neurons in response to visual stimulation at increasing contrasts, see Chance et al. 1998). Both synaptic adaptation (as in models of the phase advance in auditory nerve and V1) and intrinsic cellular properties could contribute to the observed phase-advance.

As previously noted (Rees and Moller 1983), IC responses to SAM tones are commonly nonsinusoidal in nature. A variety of nonlinearities that resemble rectification, peak clipping, phase-locking, asymmetry (possibly due to adaptation), and bimodality are seen in the period histograms over different modulation frequency ranges (see Fig. 7). The intervals between clearly visible peaks in bimodal histograms varied from 4 to 20 ms (data not shown), and probably reflect a wide variety of underlying mechanisms.

If the nonmonotonicity in pure tone RLFs is the result of a decrease in excitation, one might expect 180° phase-flips (with increasing SPL at a given modulation frequency) at low modulation frequencies (where the stimulus changes very slowly: see DISCUSSION of tMTFs above) and depths sufficiently low that the amplitude of the SAM tone traverses a SPL range that lies in the negative-slope limb of the RLF (because the response should decrease as the instantaneous SPL increases). Such 180° reversals were never seen, suggesting that the nonmonotonicity in RLFs is instead the result of an increase in inhibitory input strength (with the mean phase of the response to SAM tones reflecting the phase of the excitatory input); this conclusion is consistent with that reached from experiments using iontophoretic injection of inhibitory blockers (e.g., Fuzessery and Hall 1996).

#### *Range of BMFs*

The range of BMFs encountered in this study spans a somewhat lower range (0–140 Hz) than those reported in some of the previous studies that have measured this for different species from the rMTFs (Heil et al. 1995; Langner and Schreiner 1988; Muller-Preuss et al. 1994). In these studies, BMFs were derived from both single and multiple units; however, only one study (Langner and Schreiner 1988) differentiated the responses in the two sets. About 80% of single-unit BMFs in that study were reported to be <100 Hz, with most of the remaining <300 Hz. The slightly higher range may reflect a variety of differences between the two studies in the sampling of neurons, the species studied (cat vs. gerbil) and in the choice of BMF at 60 dB (re: threshold) versus the mean BMF (averaged across SPL and including corner frequencies) measure used in the present study. Finally, the small proportion (approximately 2%) of very high BMFs (>300 Hz) reported by Langner and Schreiner may correspond to those classified as secondary peaks in the present study.

Langner and Schreiner also observed that BMFs >300 Hz were preferentially obtained from multiple-unit recordings and suggested that this could result either from morphological differences between high BMF and low BMF neurons, or from the fact that multiple-unit responses include responses from input fibers and therefore reflect their tendency toward higher BMFs. The other two studies (Heil et al. 1995; Muller-Preuss et al. 1994 in the gerbil and squirrel monkey, respectively), reported that most BMFs lay at (and below) 160 and 128 Hz. The differences mentioned above (in the context of the Langner and Schreiner study) as well as the inclusion of multiple units may bias the range of BMFs toward a slightly higher range of values. On the other hand, a BMF distribution similar to that in the present study has been reported from the rat IC (Rees and Moller 1983); however, these investigators used a

different measure of response strength viz. the peak height of the period histogram.

Almost all BMFs in the thalamus and cortex lie below 100 Hz (Bieser and Muller-Preuss 1996; Preuss and Muller-Preuss 1990; Schreiner and Urbas 1988). The peaks in vector strength also fall within the same range. This is at least roughly consistent with the distribution of BMFs observed in the IC in the present study. Thus even though it may be potentially advantageous to create nonsynchronized responses that show BMFs at high modulation frequencies in the IC (because these responses will not be lost due to low-pass filtering at higher stages), such BMFs are rarely seen. It is, however, also possible that thalamic and cortical neurons act as band-pass filters and are therefore not very responsive to unmodulated inputs at high BMFs from IC neurons.

#### *AM response properties across carrier frequency*

Consistent with a previous study in the squirrel monkey (Muller-Preuss et al. 1994), the results of our study do not reveal any correlation of SAM response properties with the BF of the neuron in the IC. The BMF-BF relationship reported in another previous study (Langner and Schreiner 1988) is weak and is predominantly shaped by multiple-unit BMFs above 200 Hz. The increase of tMTF cutoff frequency with BF (as a result of the increase in bandwidth of the basilar membrane band-pass filter with BF) seen in auditory nerve fibers (Joris and Yin 1992) appears to be absent at the level of the IC (Fig. 12B), probably as a result of successive low-pass filtering at intervening auditory nuclei.

#### *Latency-BMF relationship*

Previous studies (Heil et al. 1995; Langner et al. 1987) have shown that the onset latency to tones at a single SPL above threshold is inversely correlated with BMF (in the range of 0–200 Hz). This finding is confirmed by our data using a slightly different analysis procedure (Fig. 11). The distribution of latencies in the present study is also similar to that in the previous two studies. The inverse correlation could be a result of the common effect of successive synaptic low-pass filtering on both measures. Alternatively, it could reflect the presence of inhibitory inputs that increase the latency and sharpen the high-frequency slope of the rMTF (thus decreasing the BMF).

#### *Topography*

A topographic map of BMFs in the IC of the cat has been reported on the basis of many closely spaced multiple-unit recordings made from two different iso-frequency laminae (Schreiner and Langner 1988). The present study was designed to allow for an extensive parametric characterization (requiring long recording times) of well-isolated single units with a range of BFs. It therefore does not provide any direct data on this issue. However, the dependence of the rMTFs on SPL suggests that a topographic map is likely to be dependent on SPL. The discrepancy between the range of BMFs observed in single and multiple units has been discussed in a preceding section. As observed by the authors of the earlier study (Schreiner and Langner 1988), because multiple-unit recordings may include responses from input fibers, it remains unclear whether the observed map reflects an organization based on the input

pattern to the IC or a topographic organization of the response properties of IC neurons.

### *Implications for perception*

Psychophysical measurements in humans and chinchillas (Salvi et al. 1982; Viemeister 1979) indicate that detection of modulation of SAM noise is possible up to modulation frequencies above 2,000 Hz. However, MTF peaks are rare above 100 Hz in both the IC (present study) and the cortex (Schreiner and Urbas 1988). MTFs measured with noise carriers from the cochlear nucleus and cortex appear to be roughly similar to those measured with tones (Eggermont 1998; Rhode 1994). The limited range of BMFs observed will therefore hamper any scheme (e.g., Langner and Schreiner 1988) based on single neurons acting as labeled lines for particular modulation frequencies by means of the peaks of their rMTFs. Further, at least in the simplest versions, such a scheme will depend on the subpopulation of neurons that remain tuned at high SPLs (because performance in modulation detection tasks remains good at high SPLs). It would also probably require them to possess a relatively invariant BMF across SPL. Thus because some neurons (at least in the IC) continue to respond at modulation frequencies above 100 Hz, information about high modulation frequencies is probably carried in some as yet unspecified manner by these response spike trains across the population of IC neurons. The secondary peaks sometimes noticed may be relevant. Also, it has recently been suggested that modulation frequencies in this high-frequency range are actually represented in the responses of neurons with frequency response areas well outside the spectrum of the presented SAM tone (Schulze and Langner 1997).

It is worth noting that most IC neurons respond with significant synchrony to modulation frequencies above 100 Hz (and below 300 Hz: Fig. 12B); tMTF cutoffs below 100 Hz are almost always because the spike rate falls to very low values. Thus considerable temporal information remains in the response in the range of modulation frequencies where most BMFs lie. Therefore the emergence of rate tuning does not necessarily preclude the possibility that information about modulation frequency is also present in a "temporal code" (for example, one based on vector strength or interspike intervals).

In human listeners, SAM tones evoke percepts of "fluctuation" below 20 Hz, "roughness" from about 20 to 300 Hz with a peak at around 70 Hz (Fastl 1990), and of pitch from about 20 to 1,000 Hz. While it is possible that the first two percepts could be encoded by schemes based either on the BMF or possibly by the phase-locked firing of neurons in the cortex (Eggermont 1998; Schulze and Langner 1997), such schemes for the third (pitch) percept run into the same problems identified while discussing the high-frequency range for the AM detection task. Further, the pitch of complex sounds is not the result of a simple determination of modulation frequency (see Cariani and Delgutte 1996 for a detailed discussion of this issue). The neural code underlying this percept at higher levels remains unclear. It is interesting in this context that there appears to be sufficient information below 50 Hz in the envelope to allow for a high degree of performance on speech recognition tasks (Shannon et al. 1995).

Human listeners can perform modulation depth discrimination tasks at all depths above the detection threshold (Lee and

Bacon 1997; Wakefield and Viemeister 1990). Because vector strengths reach close to maximum levels at very low depths in IC neurons (i.e., the measure has a limited dynamic range), some other measure like the spike rate may be a better single neuron correlate of task performance.

It is of interest to note that a recent model of the auditory processing of AM (Dau et al. 1997) is based on the presence of channels for modulation frequencies below 1,000 Hz in the auditory system. This represents a marked departure from the low-pass filter models that have been employed for some time (Viemeister 1979). Some evidence for modulation frequency specific adaptation has also been presented previously (Tansley and Suffield 1983). It would be interesting to see if the properties of perceptual channels and neuronal MTFs in the same species show any similarities.

Finally, we do not know of any current psychophysical models where the suppressive regions in the rMTF would clearly play a useful role. One might speculate that the possible sharpening of the high-frequency slopes of rMTFs might have functional implications for the selectivity of the putative psychophysical channels for modulation frequency.

We thank S. Thornton and B. Malone for participation in some of the experiments. The manuscript benefited from insightful comments by B. Malone, B. Scott, T. Lewis, R. Shapley, D. Sanes, and especially D. Tranchina and an anonymous reviewer.

This research was supported by National Institute on Deafness and Other Communication Disorders Grant DC-01767.

### REFERENCES

- ANDERSON DJ, ROSE JE, HIND JE, AND BRUGGE JF. Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects. *J Acoust Soc Am* 49: 1131–1139, 1971.
- BIESER A AND MULLER-PREUSS P. Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 108: 273–284, 1996.
- BURGER RM AND POLLAK GD. Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the inferior colliculus. *J Neurophysiol* 80: 1686–1701, 1998.
- BURKITT AN AND CLARK GM. Analysis of integrate-and-fire neurons: synchronization of synaptic input and spike output. *Neural Comput* 11: 871–901, 1999.
- CARIANI PA AND DELGUTTE B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol* 76: 1698–1716, 1996.
- CHANCE FS, NELSON SB, AND ABBOTT LF. Synaptic depression and the temporal response characteristics of V1 cells. *J Neurosci* 18: 4785–4799, 1998.
- COOPER NP, ROBERTSON D, AND YATES GK. Cochlear nerve fiber responses to amplitude-modulated stimuli: variations with spontaneous rate and other response characteristics. *J Neurophysiol* 70: 370–386, 1993.
- DAU T, KOLLMEIER B, AND KOHLRAUSCH A. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J Acoust Soc Am* 102: 2892–2905, 1997.
- EGGERMONT JJ. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J Neurophysiol* 80: 2743–2764, 1998.
- FASTL H. The hearing sensation roughness and neuronal responses to AM-tones. *Hear Res* 46: 293–295, 1990.
- FRISINA RD, SMITH RL, AND CHAMBERLAIN SC. Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement. *Hear Res* 44: 99–122, 1990.
- FUZESESSERY ZM AND HALL JC. Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. *J Neurophysiol* 76: 1059–1073, 1996.
- GOLDBERG JM AND BROWN PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32: 613–636, 1969.

- GROTJE B. Interaction of excitation and inhibition in processing of pure tone and amplitude-modulated stimuli in the medial superior olive of the mustache bat. *J Neurophysiol* 71: 706–721, 1994.
- HEIL P, SCHULZE H, AND LANGNER G. Ontogenetic development of periodicity coding in the inferior colliculus of the Mongolian gerbil. *Audit Neurosci* 1: 363–383, 1995.
- HEWITT MJ AND MEDDIS R. A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus. *J Acoust Soc Am* 95: 2145–2159, 1994.
- JORIS PX AND YIN TC. Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* 91: 215–232, 1992.
- JORIS PX AND YIN TCT. Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J Neurophysiol* 79: 253–269, 1998.
- KHANNA SM AND TEICH MC. Spectral characteristics of the responses of primary auditory-nerve fibers to amplitude-modulated signals. *Hear Res* 39: 143–157, 1989.
- KUWADA S AND BATRA R. Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit. *J Neurosci* 19: 2273–2287, 1999.
- KUWADA S, YIN TC, SYKA J, BUUNEN TJ, AND WICKESBERG RE. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. *J Neurophysiol* 51: 1306–1325, 1984.
- LANGNER G. Neuronal mechanisms for pitch analysis in the time domain. *Exp Brain Res* 44: 450–454, 1981.
- LANGNER G. Evidence for neuronal periodicity detection in the auditory system of the Guinea fowl: implications for pitch analysis in the time domain. *Exp Brain Res* 52: 333–355, 1983.
- LANGNER G, SCHREINER C, AND MERZENICH MM. Covariation of latency and temporal resolution in the inferior colliculus of the cat. *Hear Res* 31: 197–201, 1987.
- LANGNER G AND SCHREINER CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 60: 1799–1822, 1988.
- LE BEAU FE, REES A, AND MALMIERCA MS. Contribution of GABA- and glycine-mediated inhibition to the monaural temporal response properties of neurons in the inferior colliculus. *J Neurophysiol* 75: 902–919, 1996.
- LEE J AND BACON SP. Amplitude modulation depth discrimination of a sinusoidal carrier: effect of stimulus duration. *J Acoust Soc Am* 101: 3688–3693, 1997.
- MOLLER AR. Responses of units in the cochlear nucleus to sinusoidally amplitude-modulated tones. *Exp Neurol* 45: 105–117, 1974.
- MOORE DR, KOTAK VC, AND SANES DH. Commissural and lemniscal synaptic input to the gerbil inferior colliculus. *J Neurophysiol* 80: 2229–2236, 1998.
- MULLER-PREUSS P, FLACHSKAMM C, AND BIESER A. Neural encoding of amplitude modulation within the auditory midbrain of squirrel monkeys. *Hear Res* 80: 197–208, 1994.
- OLIVER DL AND HUERTA MF. Inferior and superior colliculi. In: *The Mammalian Auditory Pathway: Neuroanatomy*, edited by Webster DB, Popper AN, and Fay RR. New York: Springer-Verlag, 1991, p. 168–221.
- PRESS WH, TEUKOLSKY SA, VETTERLING WT, AND FLANNERY BP. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, MA: Cambridge Univ. Press, 1993.
- PREUSS A AND MULLER-PREUSS P. Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. *Exp Brain Res* 79: 207–211, 1990.
- REED C AND DURBECK L. Delay lines and auditory processing. In: *Comments on Theoretical Biology*. London: Gordon and Breach, 1995, p. 441–461.
- REES A AND MOLLER AR. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear Res* 10: 301–330, 1983.
- REES A AND PALMER AR. Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J Acoust Soc Am* 85: 1978–1994, 1989.
- RHODE WS. Temporal coding of 200% amplitude modulated signals in the ventral cochlear nucleus of cat. *Hear Res* 77: 43–68, 1994.
- RHODE WS AND GREENBERG S. Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 71: 1797–1825, 1994.
- ROSEN S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336: 367–373, 1992.
- RUGGERO MA. Systematic errors in indirect estimates of basilar membrane travel times. *J Acoust Soc Am* 67: 707–710, 1980.
- SALVI RJ, GIRAUDI DM, HENDERSON D, AND HAMERNIK RP. Detection of sinusoidally amplitude modulated noise by the chinchilla. *J Acoust Soc Am* 71: 424–429, 1982.
- SCHREINER CE AND LANGNER G. Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* 60: 1823–1840, 1988.
- SCHREINER CE AND URBAS JV. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear Res* 32: 49–63, 1988.
- SCHULZE H AND LANGNER G. Periodicity coding in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*): two different coding strategies for pitch and rhythm? *J Comp Physiol [A]* 181: 651–663, 1997.
- SHANNON RV, ZENG FG, KAMATH V, WYGONSKI J, AND EKELID M. Speech recognition with primarily temporal cues. *Science* 270: 303–304, 1995.
- STEPHENS MA. Tests for randomness of directions against two circular alternatives. *Am Stat Assoc J* 64: 280–289, 1969.
- TALLAL P, MILLER SL, BEDI G, BYMA G, WANG X, NAGARAJAN SS, SCHREINER C, JENKINS WM, AND MERZENICH MM. Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271: 81–84, 1996.
- TANSLEY BW AND SUFFIELD JB. Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation. *J Acoust Soc Am* 74: 765–775, 1983.
- VIEMEISTER NF. Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am* 66: 1364–1380, 1979.
- WAKEFIELD GH AND VIEMEISTER NF. Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *J Acoust Soc Am* 88: 1367–1373, 1990.
- WANG X AND SACHS MB. Transformation of temporal discharge patterns in a ventral cochlear nucleus stellate cell model: implications for physiological mechanisms. *J Neurophysiol* 73: 1600–1616, 1995.
- YANG L AND POLLAK GD. Differential response properties to amplitude modulated signals in the dorsal nucleus of the lateral lemniscus of the mustache bat and the roles of GABAergic inhibition. *J Neurophysiol* 77: 324–340, 1997.