Neural Representations of Sinusoidal Amplitude and Frequency Modulations in the Primary Auditory Cortex of Awake Primates

LI LIANG, THOMAS LU, AND XIAOQIN WANG
Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Liang, Li, Thomas Lu, and Xiaoqin Wang. Neural representations of sinusoidal amplitude and frequency modulations in the auditory cortex of awake primates. *J Neurophysiol* 87: 2237–2261, 2002; 10.1152/jn.00834.2001. We investigated neural coding of sinusoidally modulated tones (sAM and sFM) in the primary auditory cortex (A1) of awake marmoset monkeys, demonstrating that there are systematic cortical representations of embedded temporal features that are based on both average discharge rate and stimulus-synchronized discharge patterns. The rate-representation appears to be coded alongside the stimulus-synchronized discharges, such that the auditory cortex has access to both rate and temporal representations of the stimulus at high and low frequencies, respectively. Furthermore, we showed that individual auditory cortical neurons, as well as populations of neurons, have common features in their responses to both sAM and sFM stimuli. These results may explain the similarities in the perception of sAM and sFM stimuli as well as the different perceptual qualities effected by different modulation frequencies. The main findings include the following. *1*) Responses of cortical neurons to sAM and sFM stimuli in awake marmosets were generally much stronger than responses to unmodulated tones. Some neurons responded to sAM or sFM stimuli but not to pure tones. *2*) The discharge rate-based modulation transfer function typically had a band-pass shape and was centered at a preferred modulation frequency (rBMF). Population-averaged mean firing rate peaked at 16- to 32-Hz modulation frequency, indicating that the A1 was maximally excited by this frequency range of temporal modulations. *3*) Only approximately 60% of recorded units showed statistically significant discharge synchrony to the modulation waveform of sAM or sFM stimuli. The discharge synchrony-based best modulation frequency (tBMF) was typically lower than the rBMF measured from the same neuron. The distribution of rBMF over the population of neurons was approximately one octave higher than the distribution of tBMF. *4*) There was a high degree of similarity between cortical responses to sAM and sFM stimuli that was reflected in both discharge rate- or synchrony-based response measures. *5*) Inhibition appeared to be a contributing factor in limiting responses at modulation frequencies above the rBMF of a neuron. And *6*) neurons with shorter response latencies tended to have higher tBMF and maximum discharge synchrony frequency than those with longer response latencies. rBMF was not significantly correlated with the minimum response latency.

INTRODUCTION

Human speech and musical sounds contain prominent temporal modulations in both amplitude and frequency. Low-frequency (<50 Hz) modulations are important for speech perception and melody recognition, whereas modulations at higher frequencies produce other types of sensations such as pitch and roughness (Houtgast and Steeneken 1973; Rosen 1992). Amplitude and frequency modulations (AM and FM) are also important components of communication sounds of animals and are found in a wide range of species-specific vocalizations. The neural representation of amplitude- and frequency-modulated sounds begins at the auditory periphery, where auditory-nerve fibers faithfully represent both fine and coarse temporal structures of complex sounds in their temporal discharge patterns (Johnson 1980; Joris and Yin 1992; Palmer 1982). At subsequent brain stem nuclei along the ascending auditory pathway, the precision of the temporal representation degrades gradually, due to the biophysical properties of neurons along the ascending pathway and temporal integration of converging inputs from one station to the next (Blackburn and Sachs 1989; Creutzfeldt et al. 1980; de Ribaupierre et al. 1980; Frisina et al. 1990; Langner and Schreiner 1988). In a modeling study of the transformation of temporal discharge patterns from the auditory-nerve to the cochlear nucleus, Wang and Sachs (1995) showed that the reduction of phase-locking in stellate cells can result from three mechanisms: convergence of subthreshold inputs on the soma, inhibition, and the well-known dendritic low-pass filtering (Rall and Agmon-Snir 1998). These basic mechanisms may also operate at successive nuclei leading to the auditory cortex, progressively reducing the temporal limit of stimulus-synchronized responses.

It has long been known that neurons in the auditory cortex have a limited capacity to represent temporally modulated signals (Goldstein et al. 1959; deRibaupierre et al. 1972; Whitfield and Evans 1965). In contrast to subcortical neurons, neurons in the auditory cortex can only synchronize to temporally modulated signals at modulation rates of up to tens of Hertz (Eggermont 1991, 1994; Gaese and Ostwald 1995; Schreiner and Urbas 1988) compared with hundreds or thousands of Hertz subcortically (Creutzfeldt et al. 1980; Frisina et al. 1990; Joris and Yin 1992). Because most of the studies in the past three decades on this subject were conducted in anesthetized animals, with a few exceptions (Bieser and Müller-Preuss 1996; Goldstein et al. 1959; deRibaupierre et al. 1972), it has been suspected that the low temporal response rate

---

Address for reprint requests: X. Wang, Dept. of Biomedical Engineering, Johns Hopkins University School of Medicine, 720 Rutland Ave., Ross 424, Baltimore, MD 21205 (E-mail: xwang@bme.jhu.edu).

reported in the auditory cortex might partially be caused by anesthetics, which have been shown to alter temporal responses properties of the auditory cortex (Gaese and Ostwald 2001; Goldstein et al. 1959). It is therefore important to obtain measurements of cortical responses to temporally modulated signals under unanesthetized conditions, which could provide a better correlation with the perception of these signals.

While mechanisms based on stimulus-synchronized discharges have long been assumed to be the predominant means for the cortex to represent temporal modulations (see review by Langner 1992), the significance of discharge rate-based mechanisms in representing temporal features of complex sounds has gained little attention. This is perhaps due to the fact that sustained discharges are not commonly observed under anesthetized conditions. Under the awake condition, however, neurons in the auditory cortex often respond with sustained discharges throughout the entire stimulus duration (Bieser and Müller-Preuss 1996; Evans and Whitfield 1964; Lu et al. 2001a,b; Recanzone et al. 2000). Our recent study of the auditory cortex in awake primates using sequential stimuli has provided clear evidence to support a two-stage temporal processing mechanism that suggests temporal coding for slowly changing acoustic events and rate coding for rapidly changing acoustic events (Lu et al. 2001b). The present study using continuously modulated signals provides further supporting evidence for such a mechanism.

Another important issue regarding the cortical processing of temporal modulations is how cortical neurons represent similar temporal features that are introduced by different means. In the spectral domain, the notion of frequency filtering has been well established on the basis of the response area obtained from pure tones or other types of stimuli. It is not clear, however, whether a common temporal processing mechanism, or a temporal filter, is applied by cortical neurons to a variety of time-varying signals. To answer these questions, it is necessary to comprehensively test the temporal response properties of cortical neurons with a variety of temporally modulated signals. In this study, we systematically characterized cortical responses to two representative classes of temporally modulated signals, sinusoidally amplitude-modulated (sAM) and frequency-modulated (sFM) tones, in a large number of single-units in awake marmoset monkeys (*Callithrix jacchus*), a highly vocal primate species (Wang 2000). Preliminary observations from the present study were presented at two conferences (Liang et al. 1999; Wang et al. 2001).

METHODS

Animal preparation and recording procedures

Details on animal preparation and recording procedures were described in a previous study (Lu et al. 2001a) and are only briefly described here. Marmosets were adapted to sit quietly during recording sessions in an apparatus specially designed for this species. The auditory cortex was accessed laterally using a single tungsten microelectrode of impedance typically ranging from 2 to 5 MΩ at 1 kHz (A-M Systems) through a small hole (diameter, ~1.0 mm) in the skull. Only one opening in the skull existed at any given time during the recording sessions. Each hole was sealed by dental cement after several days of recordings. Necessary steps were taken to ensure sterility during all recording sessions. Daily recording sessions (3–5 h) were carried out for several months in each animal. The advantage of this procedure was that it only left a very small portion of the cortex

exposed, which greatly increased the recording stability, avoided excess tissue growth and reduced the chance of infections through the opening. All recording sessions were conducted in a double-walled, sound-proof chamber (IAC-1024). The interior of the chamber was covered by 3-in acoustic absorption foam (Sonex, Illbruck). The experimental procedures were approved by the Animal Care and Use Committee at The Johns Hopkins University.

Densely positioned recording holes were made covering the auditory cortex. The data presented were mainly obtained from the primary auditory cortex (A1) and may include a few neurons from the immediately adjacent areas that responded to sAM and/or sFM stimuli. The location of A1 was determined by its tonotopic organization, its relationship to the lateral belt area (which was more responsive to noises than tones), and by its response properties (e.g., highly responsive to tonal stimuli). Electrode penetrations, perpendicular to the cortical surface, were made within each recording hole under visual guidance via an operating microscope. This gave good control and estimates of recording depths. Single-units were encountered at all cortical layers, but the majority of the recorded units was from upper layers, judging by the depths and response characteristics. On average, one to three well-isolated single-units were studied in each daily session. A representative example of raw recordings is shown in Fig. 1. Signal-to-noise ratio was typically >10:1 in our recordings. Spike waveforms were filtered, digitized, and detected using a template-matching discriminator (MSD, Alpha-Omega Engineering) and were closely and constantly monitored by an experimenter as the recording proceeded. The template matching method prevented any unwanted noises (e.g., due to the animal's movement) from triggering false spikes.

Acoustic stimuli

Two types of temporally modulated sounds were used as the experimental stimuli in this study: sAM and sFM sounds. For sAM stimuli, the carrier frequency, set at a unit's characteristic frequency (CF), was held constant while its amplitude was modulated by a sinusoid. For sFM sounds, the amplitude remained constant while the carrier frequency, centered at a unit's CF, was sinusoidally modulated. Acoustic stimuli were delivered in free-field through a loudspeaker located ~1 meter in front of the animal. Frequency tuning was obtained using a series of randomly presented tone bursts (50–100 ms in duration), from which the CF of a unit was determined as the frequency at which the strongest discharge rate was evoked. A rate-level function was obtained at CF. Most units in the awake auditory cortex exhibited nonmonotonic discharge rate versus sound level functions (Pfingst and O'Connor 1981; Wang et al. 1999), for which a preferred sound level can be determined. After these initial characterizations, multiple repetitions of sAM and sFM stimuli were delivered at the preferred sound level for nonmonotonic units (or 30 dB above threshold if otherwise). In a subset of units, sAM and/or sFM stimuli were tested at multiple sound levels. Several stimulus parameters were varied to probe neural responses. For every unit included in the analysis, modulation frequency was typically varied between 1 and 512 Hz in a base-2 logarithmic scale; finer steps were used in testing some units. The modulation depth of sAM stimuli was generally set at 100%; in a subset of sampled units, a range of depth (0–100%) was tested. The modulation depth of sFM stimuli was usually set at an optimal (in terms of maximum firing rate) or a nearly optimal depth centered at a unit's CF. Multiple FM depths were tested in some units. Unlike sAM stimuli, the firing rate resulting from a sFM stimulus was not monotonically related to modulation depth. The optimal depth, at which the maximum firing rate was achieved, varied from unit to unit. It was therefore not feasible to choose a fixed FM modulation depth for all units. Using the optimal FM depth for a unit allowed measurement of modulation selectivity to be made at its maximum discharge level, thus increasing the robustness of the measurement. Because firing rates of a unit at a given modulation frequency generally increased with increasing AM depth for a sAM
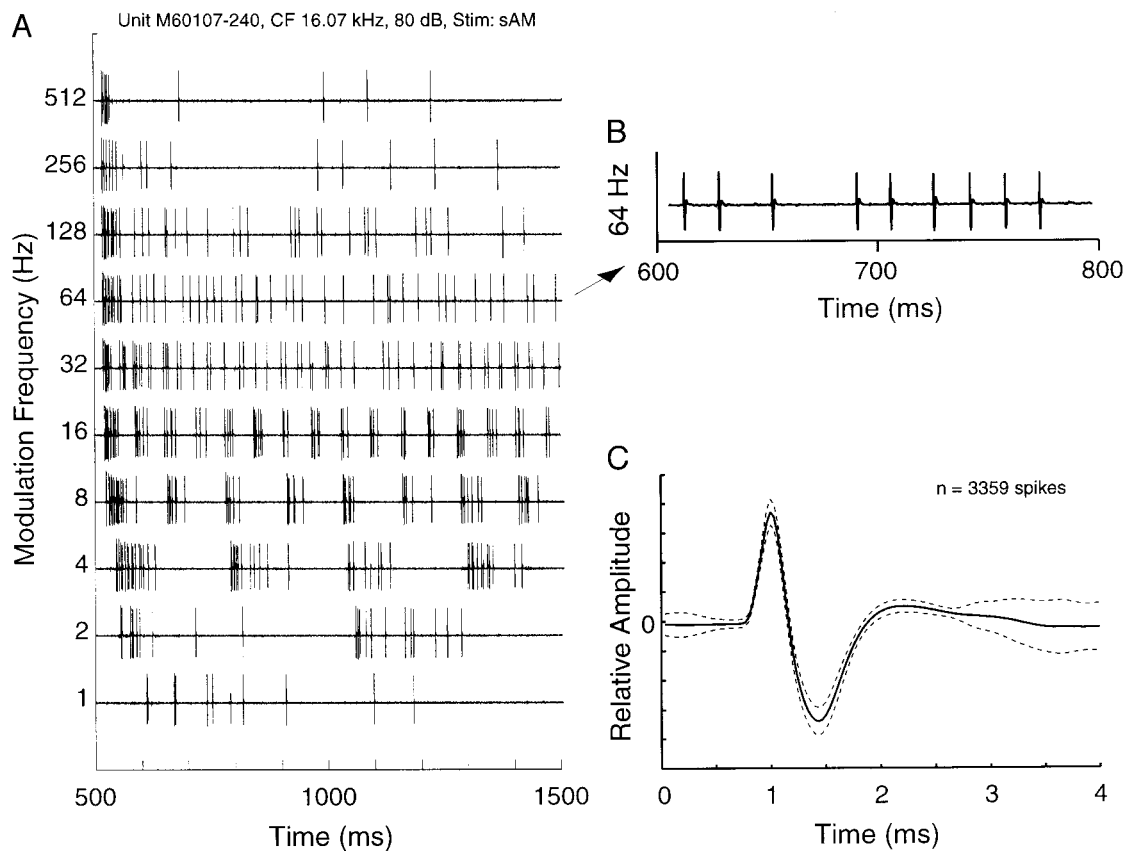
FIG. 1. Examples of digitized traces of extracellular recordings from the auditory cortex of an awake marmoset in response to sinusoidally amplitude-modulated (sAM) stimuli. Action potentials from a single-unit are present in the recordings. A complete display of this unit's responses to sAM and sinusoidally frequency-modulated (sFM) stimuli is given in Fig. 2B. A: digitized traces are ordered vertically according to modulation frequency (from 1 to 512 Hz as indicated along ordinate). Stimulus duration was 1,000 ms (onset at 500 ms). One repetition for each modulation frequency is shown. Signal-to-noise ratio was typically >10:1 in our recordings. B: a portion of the response at 64-Hz modulation frequency is shown in an enlarged view. C: average waveform of all action potentials (n = 3,359) recorded during the presentation of the entire set of sAM stimuli from this unit (10 repetitions at each of 10 modulation frequencies). —, the mean value; - - - , 1 SD away from the mean. Sampling rate for the data shown in this figure was 30 kHz. Quantitative analyses of responses recorded from this unit are subsequently shown in Figs. 2–4B and 7B.

stimulus, it was possible to used a fixed AM depth (100%) to test all units. The duration of each sAM and sFM stimulus was 1,000 ms. Neural activities prior to and following stimulus presentation were also recorded to estimate spontaneous discharges and to reveal any long-lasting effects. Ten to 20 repetitions of a sAM or sFM stimulus were presented at each modulation frequency at a given sound level. Stimuli of all modulation frequencies were presented randomly. Inter-stimulus intervals were >1 s. Stimuli were synthesized at 100-kHz sampling rate and low-pass filtered at 50 kHz. The spike times were digitized at 50-kHz sampling rate.

Data analysis

The results reported were based on 211 single-units recorded from the left hemispheres of three awake marmoset monkeys. Responses to sAM and sFM stimuli were recorded in 200 and 142 units, respectively. Some units responded to both types of stimuli and others responded to either sAM or sFM stimuli. All units were recorded indiscriminately, provided they could be driven by either sAM or sFM stimuli. Given the diversity of cortical responses in the awake preparation, we did not expect all recorded units to respond to both types of stimuli. Discharge rates of the majority of units varied with changing modulation frequency. We separated sAM and sFM responses, respectively, into two groups in the analyses according to the characteristics of the discharge rate versus modulation frequency profile of

a unit. A rate profile was considered having a "band-pass" shape if its peak value was higher than values at both the lower and higher frequency sides. Some units with band-pass rate profiles also exhibited increased discharge rates with increasing modulation frequency at frequencies much higher than the frequency corresponding to the peak of the rate profile. This portion of responses was not included in the analysis of rate-based modulation selectivity because they were more likely to be influenced by spectral effects at such high modulation frequencies. Responses with band-pass rate profiles were further screened on the basis of a d' value, $d' = |\mu_{Rd} - \mu_{Rs}|/\sigma_{Rs}$, where μ_{Rd} is the mean discharge rate during a stimulus presentation, μ_{Rs} is the mean spontaneous discharge rate, and σ_{Rs} is the SD of the spontaneous discharge rate. Values of d' were calculated for responses to each modulation frequency tested. If a unit had a band-pass rate profile and a maximum d' value ≥ 1.0, it was classified into the band-pass (BP) group. The rest of the units were referred to as the nonband-pass (non-BP) group, which included units whose maximum d' values were <1.0 (indicating weak responses to sAM or sFM stimuli) as well as those units whose rate profiles did not exhibit "band-pass" shapes. A unit belonged to either the BP or non-BP group. This classification process was separately applied to sAM and sFM responses. For sAM responses, 146/200 (73%) units belonged to the BP group and 54/200 (27%) units belonged to the non-BP group. For sFM responses, the numbers were 93/142 (65%, BP group) and 49/142 (35%, non-BP group), respectively. Units in the BP group responded to the *change*

in modulation frequency by their firing rates and were analyzed for discharge rate-based modulation selectivity. Units in both BP and non-BP groups were analyzed for discharge synchrony-based modulation selectivity. Population averages were computed for each group of units as well as for all the units.

Unless specified, average discharge rates were calculated over a window including the stimulus duration and 100 ms after stimulus offset: (t_{onset}, t_{offset} + 100 ms), where t_{onset} and t_{offset} are stimulus onset and offset times. Spontaneous discharge rates were estimated from activities prior to stimulus onset (500 ms) and subtracted from raw discharge rates. Several response measures, described in the following text, were used to quantify the modulation selectivity of cortical neurons. Comparison between response measures resulting from sAM and sFM stimuli were made between populations of units that responded to each stimulus type as well as on a unit-by-unit basis in individual units in which both measures could be obtained. Statistical comparisons between distributions of response measures were made using the Wilcoxon rank-sum test (Rice 1988). Unit-by-unit comparisons between response measures recorded from the same unit were made using the paired t-test. $P < 0.01$ was considered statistically significant for these analyses.

RATE MODULATION TRANSFER FUNCTION (rMTF). The relationship between average discharge rate and modulation frequency is referred to as the *discharge rate-based modulation transfer function* (rMTF). The *discharge rate-based best modulation frequency* (rBMF) was calculated in the following steps for each unit belonging to the BP-group. For units with peaks in their rMTFs that consisted of two or more points, an estimate of rBMF was first obtained as the modulation frequency corresponding to the largest discharge rate of a rMTF. The rBMF was then calculated by weighting those modulation frequencies that were continuous and adjacent to the estimated rBMF and whose discharge rates were not significantly different from the estimated rBMF ($P > 0.05$, Wilcoxon rank-sum test). A geometric mean was used to average the modulation frequencies by their discharge rates. The rBMF would be equal to the estimated rBMF if discharge rates at other frequencies were all significantly different from that at the estimated rBMF. This method has an advantage over simply assigning the rBMF to one of tested modulation frequencies corresponding to the largest discharge rate, as commonly used in most previous studies. For example, a rMTF that had a broad peak centered on two modulation frequencies with similar discharge rates would have a calculated rBMF near their geometric mean but weighted closer to the modulation frequency that produced a stronger discharge.

To calculate the *half-height bandwidth* (BW) of a rMTF, the initial estimate of the rBMF and the corresponding discharge rate were used as the reference. Two points on the rMTF curve, each on one side of the estimated rBMF, that were nearest to half the discharge rate at the estimated rBMF were interpolated linearly from the tested modulation frequencies above and below the half-discharge rate. The distance between the modulation frequencies corresponding to these two points was the BW of a rMTF. A Q-measure, defined as rBMF/BW, was used to quantify the sharpness of tuning of a rMTF.

SYNCHRONY MODULATION TRANSFER FUNCTION (tMTF). Stimulus-synchronized discharges were characterized first by the vector strength (VS) (Goldberg and Brown 1969) and then converted to the Rayleigh statistics ($2n\text{VS}^2$, where n is the total number of spikes) (Mardia and Jupp 2000) to assess their statistical significance. VS was calculated with a time window beginning 100 ms after stimulus onset to the end of the stimulus. The values of the Rayleigh measure >13.8 were considered as statistically significant ($P < 0.001$) (Mardia and Jupp 2000). The relationship between Rayleigh statistics and modulation frequency was referred to as the *temporal modulation transfer function* (tMTF). In some units, tMTF had a band-pass shape. A response measure called *discharge synchrony-based best modulation frequency* (tBMF) was calculated from each tMTF. For units whose tMTFs consisted of two or more significant values (Rayleigh statistic > 13.8), the tBMF was obtained by weighting

modulation frequencies with significant VS that were adjacent, continuous and surrounding the modulation frequency corresponding to the maximum VS. A geometric mean was used to average the modulation frequencies by their VS. To calculate *maximum synchronization frequency* (f_{max}), we first determined the highest modulation frequency at which significant discharge synchrony was found. A linear interpolation was made between this frequency and the adjacent, higher, tested modulation frequency with nonsignificant Rayleigh statistics. The modulation frequency where the interpolated Rayleigh statistic line crossed 13.8 was taken to be f_{max}.

CALCULATION OF THE DOUBLING OF SYNCHRONIZATION FREQUENCY. A subset of units exhibited discharge patterns that were synchronized to twice the modulation frequency ($2f_m$) under certain stimulus conditions. For such units, tMTFs based on Rayleigh statistics were calculated at f_m and $2f_m$, respectively [i.e., $2n(\text{VS}_{fm})^2$ and $2n(\text{VS}_{2fm})^2$]. A unit was considered having the doubling of synchronization frequency if the peak of the tMTF($2f_m$) was higher than the magnitude of the tMTF(f_m) at the corresponding modulation frequency. For this type of unit, f_{max} based on both tMTF(f_m) and tMTF($2f_m$) were computed in the same manner as described in the preceding text.

MINIMUM RESPONSE LATENCY. Minimum response latency to sAM or sFM stimuli was determined, respectively, on the basis of a composite PSTH of a unit's responses at all tested modulation frequencies. A cumulative post-stimulus histogram (PSTH) was then constructed by integrating the PSTH over time. The time after stimulus onset at which the spike count in a bin of the cumulative PSTH exceeded twice the largest SD of spike counts prior to stimulus onset was calculated as the *minimum response latency*. The binwidth used in the calculation was 1 ms. The minimum response latency defined here could be different from the first spike latency measured from CF tones in each unit. We used this latency measure instead of the first spike latency because it was a more direct indicator of onset timing of a neuron in response to sAM or sFM stimuli. The first spike latency was not always available as some neurons did not respond well to unmodulated CF tones.

RESULTS

General observations

The majority of the sampled neurons responded to CF tones as well sAM or sFM stimuli. However, neurons generally responded more strongly to sAM and sFM stimuli, often with sustained firing, than to tones as judged by number of spikes evoked over the duration of the sAM and sFM stimuli (1 s). Some of the neurons could only be driven by sAM or sFM stimuli with proper parameters (e.g., modulation frequency and depth). Representative examples of responses to sAM and sFM stimuli are shown in Fig. 2. Overall discharge patterns can be seen from the dot raster and PSTH. Temporal discharge patterns are further illustrated by period histograms computed from the same group of units in Fig. 3. These examples show that responses of units generally varied as a function of modulation frequency. They also show that some, but not all, recorded units exhibited stimulus-synchronized discharges at low modulation frequencies that gradually disappeared with increasing modulation frequency (Fig. 2, A–C). Responses to sAM and sFM stimuli often diminished at high modulation frequencies for the frequency range tested (Fig. 2Db). Not all units responded at low modulation frequencies (Fig. 2, Da and Ea). In general, sustained discharges were limited to a narrower range of modulation frequencies than onset discharges (Fig. 2, D and E). These and other response characteristics are quantitatively analyzed in the following sections.
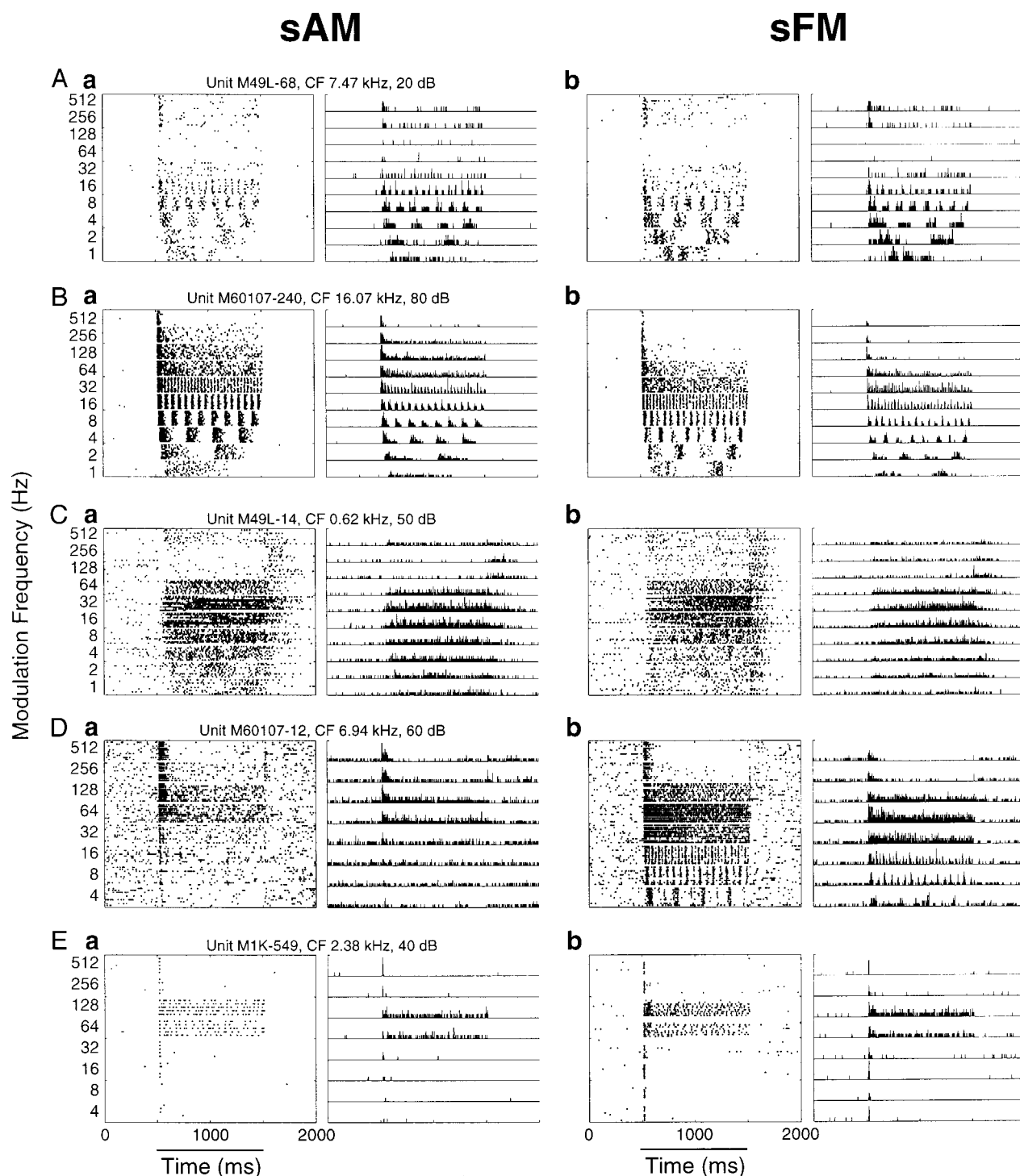
FIG. 2.   Examples of cortical responses to sAM and sFM stimuli from 5 representative single-units. *A–E*: each row shows responses to both sAM (*a*) and sFM (*b*) stimuli recorded from each unit in the form of dot raster (*left*) and poststimulus time histogram (PSTH; *right*). The displays are arranged according to modulation frequency along the ordinate. Stimulus duration was 1,000 ms (onset at 500 ms), as indicated by a horizontal bar below time axis (*E*). Stimulus parameters for each example are as follows f_c, carrier frequency; d_{sAM}, sAM modulation depth; d_{sFM}, sFM modulation depth. *A*: f_c = 7.47 kHz, 20 dB SPL, d_{sAM} = 100%, d_{sFM} = 256 Hz. *B*: f_c = 16.07 kHz, 80 dB SPL, d_{sAM} = 100%, d_{sFM} = 1,024 Hz. *C*: f_c = 0.62 kHz, 50 dB SPL, d_{sAM} = 100%, d_{sFM} = 32 Hz. *D*: f_c = 6.94 kHz, 60 dB SPL, d_{sAM} = 100%, d_{sFM} = 3566 Hz. *E*: f_c = 2.38 kHz, 40 dB SPL, d_{sAM} = 100%, d_{sFM} = 362 Hz.

Discharge rate-based modulation frequency selectivity

PROPERTY OF INDIVIDUAL NEURONS.   The majority of recorded units displayed selectivity for a particular modulation frequency when measured by average discharge rate. Figure 2

shows responses to both sAM (*a*) and sFM (*b*) stimuli recorded from the same units across a range of modulation frequencies. rMTF from all five units shown in Fig. 2 are plotted in Fig. 4. For example, the unit shown in Fig. 2*C* responded most
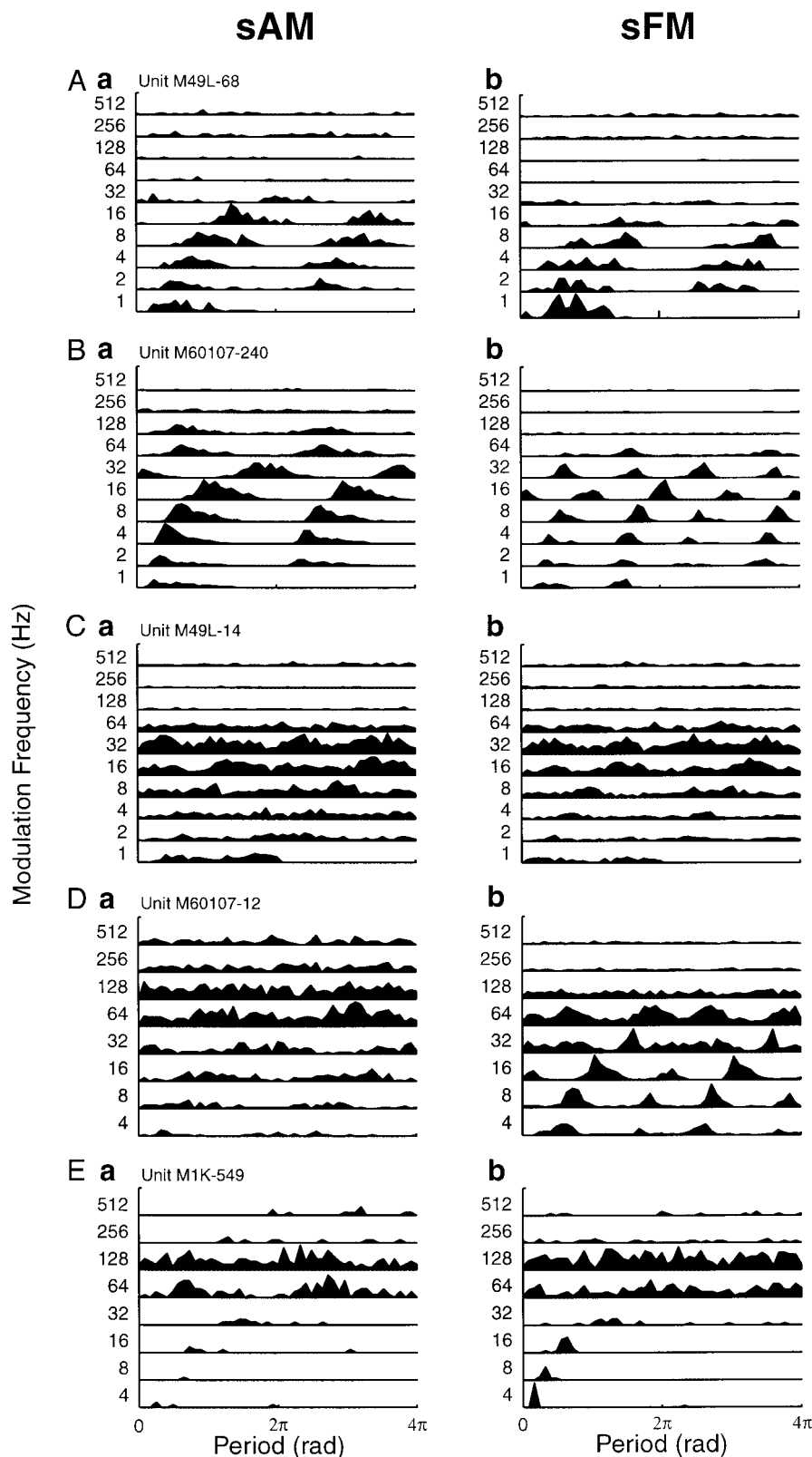
FIG. 3. Period histograms calculated from the responses shown in Fig. 2. *A–E* correspond to those in Fig. 2, *A–E,* respectively. Responses to sAM and sFM are shown in *a* and *b*, respectively. Two stimulus periods are shown in each histogram. The periods are shown in units of radians.

strongly near 32-Hz modulation frequency. PSTHs in Fig. 2*C* show sustained firing at a modulation frequency of 32 Hz for both sAM (Fig. 2*Ca*) and sFM (Fig. 2*Cb*) stimuli. rMTFs produced by sAM and sFM stimuli had similar band-pass shapes and peaked at a modulation frequency of 32 Hz (Fig.

4*C*). The peak in a rMTF is conventionally referred to as the rBMF. In this study, we used a quantitative method to calculate the rBMF instead of using a particular tested modulation frequency (see METHODS). Other examples in Fig. 2 illustrate the typical range of modulation frequency selectivity observed in

the recorded units. A prominent feature in these examples, representative of our large samples, is the sustained firing for the entire stimulus duration at modulation frequencies near rBMF (Fig. 2, *D* and *E*). Neurons generally responded more weakly at modulation frequencies lower than rBMF. In some cases, there were no responses or only brief onset responses at these lower modulation frequencies (e.g., Fig. 2*E*). The disappearance of sustained discharges at modulation frequencies higher than rBMF was commonly observed (e.g., Fig. 2, *B–E*).
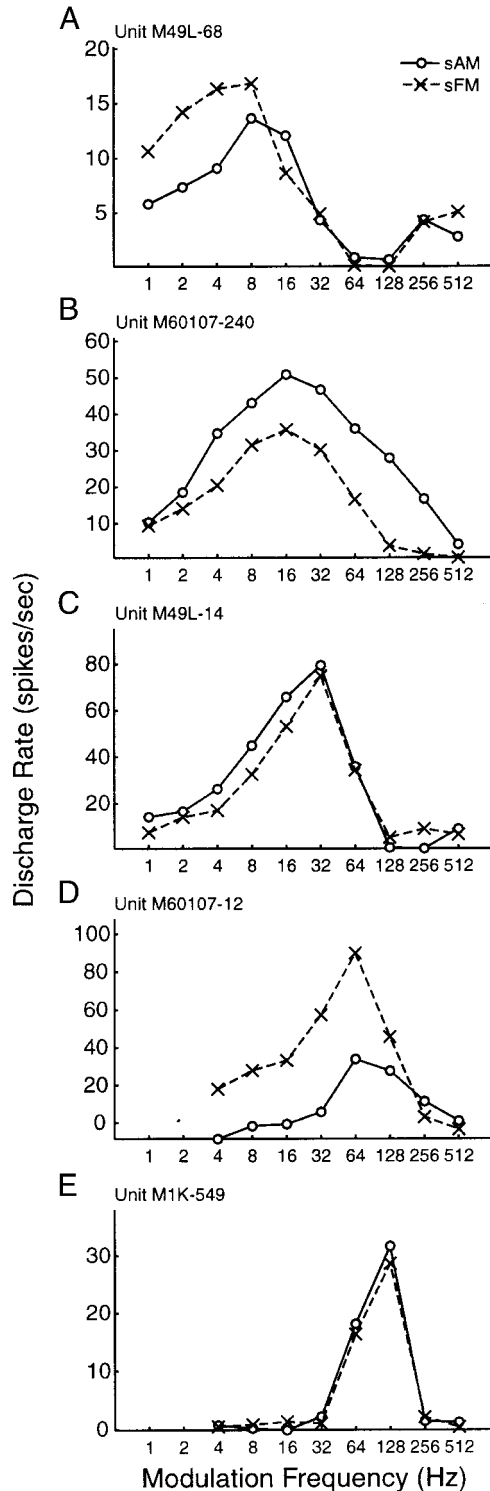
The lack of responses at high modulation frequencies appeared to result from inhibition in many cases (e.g., Fig. 2, *C* and *D*). In general, rMTFs derived from responses of a unit to sAM and sFM stimuli had similar shapes and closely matched rBMF.

POPULATION PROPERTIES.   Using the procedures described in METHODS, we have computed rBMF from the units that responded reliably to sAM and/or sFM stimuli (see METHODS). Figure 5*A* shows distributions of $rBMF_{sAM}$ and $rBMF_{sFM}$, respectively. Both distributions are centered between 16 and 32 Hz ($rBMF_{sAM}$: median = 22.6 Hz, $rBMF_{sFM}$: median = 18.1 Hz, see Table 1) and are statistically indistinguishable (Wilcoxon rank-sum test, $P = 0.1$). Among the 211 single-units that we studied, $rBMF_{sAM}$ could be determined in 146 units (69%), whereas as $rBMF_{sFM}$ could be determined in 93 units (44%). Figure 5*B* shows the relationship between $rBMF_{sAM}$ and $rBMF_{sFM}$ in 76 units where both were determined in the same unit. $rBMF_{sAM}$ and $rBMF_{sFM}$ were highly correlated (correlation coefficient $r = 0.7$, Table 1). A paired *t*-test showed that there was no significant difference ($P = 0.02$) between $rBMF_{sAM}$ and $rBMF_{sFM}$ when compared on a unit-by-unit basis. A closer examination revealed that units in the upper 50th percentile of $rBMF_{sAM}$ ($rBMF_{sAM} >$ median $rBMF_{sAM} = 20.7$ Hz) appeared to have significantly higher $rBMF_{sAM}$ than $rBMF_{sFM}$ values (paired *t*-test, $P < 0.01$). There was no statistically significant difference (paired *t*-test, $P = 0.09$) between $rBMF_{sAM}$ and $rBMF_{sFM}$ for the units in the lower 50th percentile of $rBMF_{sAM}$ (Table 1). The distribution of the difference between $rBMF_{sAM}$ and $rBMF_{sFM}$ pairs is shown in Fig. 5*C*. The close match between $rBMF_{sAM}$ and $rBMF_{sFM}$ was found in a substantial proportion of this population of units. The median of the distribution was 0.46 octaves, with 50 of 76 units (66%) having closely matched $rBMF_{sAM}$ and $rBMF_{sFM}$ (differences within 1.0 octave). These data show that there was a great degree of similarity in the responses to sAM and sFM stimuli, as reflected in mean firing rate, both at the level of single neurons and the level of populations of neurons in the auditory cortex. In general, the match between the $rBMF_{sAM}$ and $rBMF_{sFM}$ was independent of whether discharges were synchronized to the modulation waveform or not.

BANDWIDTH OF rMTF.   In Fig. 6 we analyzed and compared half-height BW and sharpness of tuning (*Q*) of rMTF for both sAM and sFM stimuli (see METHODS). The distributions of BW across populations of the neurons were similar for both types of stimuli (Fig. 6*A*) and were not statistically different (Wilcoxon

FIG. 4.   Examples of discharge rate-based modulation transfer functions (rMTF). Average discharge rates were calculated over a period including the stimulus duration and 100 ms after stimulus offset, with spontaneous discharge rates subtracted (see METHODS). Units in *A–E* correspond to those shown in Fig. 2, *A–E*, respectively. For each unit, rMTFs produced by sAM (○—○) and sFM (×—×) stimuli are shown. The best modulation frequencies (rBMF) calculated from rMTF (see METHODS) for the 5 examples shown are as follows ($rBMF_{sAM}$, $rBMF_{sFM}$). *A*: 11.07, 4.16 Hz. *B*: 16.0, 16.0 Hz. *C*: 23.38, 32.0 Hz. *D*: 87.36, 64.0 Hz. *E*: 128.0, 128.0 Hz. The half-height bandwidth (BW) and its corresponding low- and high-frequency boundaries (see METHODS) are listed in the following order: BW_{sAM} (low-frequency, high-frequency), BW_{sFM} (low-frequency, high-frequency). *A*: 25.19 (1.65, 26.84) Hz, none (none, 16.98) Hz. *B*: 153.23 (2.85, 156.08) Hz, 57.71 (3.19,60.9) Hz. *C*: 54.39 (6.88, 61.27) Hz, 51.74 (9.87, 61.61) Hz. *D*: 166.94 (44.77, 211.71) Hz, 106.02 (23.86, 129.88) Hz. *E*: 138.12 (59.07 197.19) Hz, 137.34 (59.7 197.04) Hz. The *Q* measure, defined as rBMF/BW, is listed in the following order: Q_{sAM}, Q_{sFM}. *A*: 0.44, none. *B*: 0.1, 0.28. *C*: 0.43, 0.62. *D*: 0.52, 0.6. *E*: 0.93, 0.93.
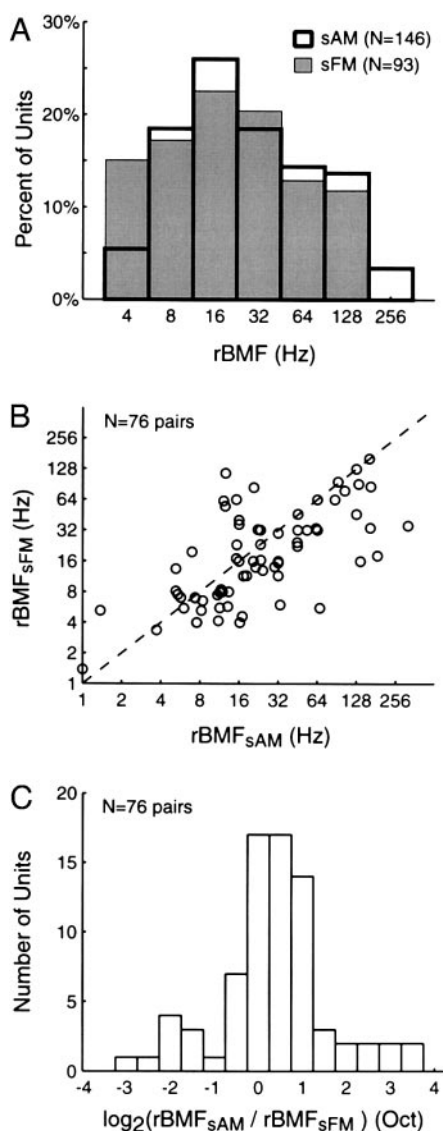
FIG. 5. Population properties for discharge rBMF selectivity. *A*: overlapping histograms showing distributions of rBMF derived from sAM (□) and sFM (■) stimuli, respectively. The binwidths of the histograms are on a base-2 logarithmic scale. The distributions of $rBMF_{sAM}$ and $rBMF_{sFM}$ are not statistically different from each other (Wilcoxon rank-sum test, $P = 0.1$). The means of the 2 distributions are 25.9 Hz ($rBMF_{sAM}$) and 19.2 Hz ($rBMF_{sFM}$) on the base-2 logarithmic scale and 48.8 Hz ($rBMF_{sAM}$) and 35.1 Hz ($rBMF_{sFM}$) on the linear scale, respectively (Table 1, columns 1–3). The medians of the 2 distributions are 22.6 Hz ($rBMF_{sAM}$) and 18.1 Hz ($rBMF_{sFM}$), respectively. *B* and *C*: unit-by-unit comparison between $rBMF_{sFM}$ and $rBMF_{sAM}$ recorded from the same unit (Table 1, column 4). This group of 76 units was a subset of the units analyzed in *A*. $rBMF_{sAM}$ and $rBMF_{sFM}$ were not statistically different from each other in the unit-by-unit comparison (paired *t*-test, $P = 0.02$). The correlation coefficient (r) between $rBMF_{sAM}$ and $rBMF_{sFM}$ (*B*) was 0.7. The difference between $rBMF_{sAM}$ and $rBMF_{sFM}$ was calculated in octaves as $\log_2(rBMF_{sAM}/rBMF_{sFM})$ (*C*, median = 0.46 octaves). - - -, slope of 1.0.

rank-sum test, $P = 0.16$). The distributions of Q values (Fig. 6*C*) were also similar between sAM and sFM stimuli (Wilcoxon rank-sum test, $P = 0.77$). The medians of BW distributions were between 32 and 64 Hz (BW_{sAM}: 53.9 Hz, BW_{sFM}: 49.3 Hz, see Table 1), approximately one octave greater than the medians of rBMF (Fig. 5), which resulted in Q distributions centered around 0.5 octaves (Q_{sAM}: 0.45, Q_{sFM}: 0.47). In a

subpopulation of units, BW and Q could be measured for both sAM and sFM stimuli in the same units. Figure 6, *B* and *D*, shows that, when compared on a unit-by-unit basis, the two measures of the tuning width of rMTF were not statistically different (paired *t*-test, BW: $P = 0.27$, Q: $P = 0.1$) between sAM and sFM stimuli (Table 1). These results showed that in addition to the similarity in rBMF, there was also a similarity in the sharpness of tuning of rMTF produced by sAM and sFM stimuli, both at the level of single neurons and across populations of neurons.

Discharge synchrony-based modulation frequency selectivity

PROPERTY OF INDIVIDUAL NEURONS. The examples given in Figs. 2 and 3 also show that discharges of cortical neurons in response to sAM and sFM stimuli could exhibit stimulus-synchronized temporal patterns. Period histograms shown in Fig. 3, *A–E*, corresponding to the units shown in Fig. 2, *A–E*, further illustrate temporal discharge patterns evoked by the modulated sounds. Discharges synchronized to modulation waveform should show peaks in both of the two periods plotted (Fig. 3). Discharges registered in the first but not the second period of a histogram indicate that they were synchronized to stimulus onset but not to the modulation waveform. The unit in Fig. 3*B* responded to sAM stimuli with well-synchronized discharges at modulation frequencies ≤128 Hz as can been seen from the period histogram. The phase delay increased with increasing modulation frequency. At a 32-Hz modulation frequency, discharges from preceding period began to appear (Fig. 3*Ba*). The temporal discharge patterns in response to sFM stimuli (Fig. 3*Bb*) in the same unit differed markedly from those to sAM stimuli (Fig. 3*Ba*) in that there were two clusters of firings within each modulation period. This was because both the upward and downward trajectory of the modulation waveform excited this unit. In general, when sAM stimuli were used, discharges could be synchronized at a rate approximately equal to the modulation frequency, whereas for sFM stimuli, response synchronization could occur at a rate twice as large as the modulation frequency. Moreover, stimulus-induced synchronization was sometimes produced by one type of the modulated sounds but not by another type in an individual unit. For example, the unit in Figs. 2*D* and 3*D* responded with synchronized discharges to sFM but not sAM stimuli.

We used the vector strength (VS) to quantify stimulus-synchronized firing patterns and Rayleigh statistics to assess the statistical significance (see METHODS) because low firing rates undermine the interpretation of the VS measure. In Fig. 7, stimulus-synchronized discharges were quantified for the units described in Figs. 2 and 3 in the form of tMTF (see METHODS). Significant discharge synchronization was found in most, but not all, recorded units. The unit shown in Figs. 2*B*, 3*B*, and 7*B* was an example with synchronized discharges. Stimulus-synchronized discharges were present in this unit at modulation frequencies ≤128 Hz for sAM stimuli and were strongest at 16 Hz (Fig. 7*B*). This peak in tMTF has traditionally been referred to as tBMF. tBMF was quantitatively determined in this study using a weighting method (see METHODS). Because this unit apparently responded to both the upward and downward trajectory of the sFM stimuli, the Rayleigh statistics calculated based on the modulation frequency of the sFM stimuli had

TABLE 1. *Measures of response to sAM and sFM stimuli*

	sAM Response	sFM Response	sAM vs. sFM Population Comparison	sAM vs. sFM Unit-by-Unit Comparison
rBMF, Hz	$\mu \pm$ std (log): 25.9 [8.3, 80.5] $\mu \pm$ std (linear): 48.8 $\pm$ 63.8 median [25%, 75%]: 22.6 [11.4, 64.0] $n=146$	$\mu \pm$ std (log): 19.2 [6.3, 61.2] $\mu \pm$ std (linear): 35.1 $\pm$ 38.0 median [25%, 75%]: 18.1 [7.7, 44.9] $n=93$	$P = 0.10$	$r = 0.70$ $P = 0.02$ (all units) $P = 0.09$ (lower 50% rBMF$_{sAM}$, $n = 38$) $P < 0.01$ (upper 50% rBMF$_{sAM}$, $n = 38$) median rBMF$_{sAM}$: 20.7 Hz median rBMF$_{sFM}$: 16.2 Hz $n=76$
BW, Hz	$\mu \pm$ std (log): 64.0 [20.1, 203.4] $\mu \pm$ std (linear): 125.1 $\pm$ 171.6 median [25%, 75%]: 53.9 [25.4, 155.6] $n=146$	$\mu \pm$ std (log): 49.1 [14.7, 164.3] $\mu \pm$ std (linear): 98.2 $\pm$ 130.8 median [25%, 75%]: 49.3 [21.4, 106.0] $n=93$	$P = 0.16$	$r = 0.49$ $P = 0.27$ median BW$_{sAM}$: 49.6 Hz median BW$_{sFM}$: 46.8 Hz $n=76$
Q	$\mu \pm$ std: 0.53 $\pm$ 0.39 median [25%, 75%]: 0.45 [0.27, 0.68] $n=146$	$\mu \pm$ std: 0.64 $\pm$ 0.66 median [25%, 75%]: 0.47 [0.28, 0.69] $n=93$	$P = 0.77$	$r = 0.31$ $P = 0.10$ median Q$_{sAM}$: 0.43 median Q$_{sFM}$: 0.47 $n=76$
tBMF, Hz	$\mu \pm$ std (log): 9.7 [3.9, 24.1] $\mu \pm$ std (linear): 15.6 $\pm$ 21.4 median [25%, 75%]: 9.6 [5.4, 16.0] $n=134$	$\mu \pm$ std (log): 9.2 $\pm$ [3.7, 23.4] $\mu \pm$ std (linear): 14.2 $\pm$ 17.6 median [25%, 75%]: 10.0 [4.0, 16.0] $n=75$	$P = 0.82$	$r = 0.50$ $P = 0.95$ median tBMF$_{sAM}$: 10.8 Hz median tBMF$_{sFM}$: 10.9 Hz $n=59$
f_{max}, Hz	$\mu \pm$ std (log): 34.2 [11.3, 103.1] $\mu \pm$ std (linear): 58.9 $\pm$ 66.2 median [25%, 75%]: 34.2 [17.0, 63.6] $n=134$	$\mu \pm$ std (log): 32.9 [10.4, 104.4] $\mu \pm$ std (linear): 57.4 $\pm$ 61.7 median [25%, 75%]: 39.4 [13.4, 79.4] $n=75$	$P = 0.97$	$r = 0.70$ $P = 0.11$ median f_{max}(sAM): 47.1 Hz median f_{max}(sFM): 40.1 Hz $n=59$

Statistics shown are based on the data presented in Fig. 5 [rate-based best modulation frequency (rBMF)], Fig. 6 [bandwidth (BW, Q)], Fig. 8 [synchrony-based best modulation frequency (tBMF)], and Fig. 9 (f_{max}), respectively. Values are mean $\pm$ SD of a distribution calculated on the base-2 logarithmic scale and converted to the linear scale using the formula: $2^{(\log-\mu)}$ and $[2^{(\log-\mu-\log-std)}, 2^{(\log-\mu+\log-std)}]$, where log-$\mu$ and log-std are the mean $\pm$ SD calculated on the base-2 logarithmic scale, respectively. r is the correlation coefficient based on the log2-scale for rBMF, tBMF and f_{max}, and on the linear scale for BW and Q. n is number of units. P is the Wilcoxon rank-sum test score for comparison between responses to sinusoidally amplitude-modulated (sAM) and sinusoidally frequency-modulated (sFM) stimuli (column 3) and the paired *t*-test score for sAM versus sFM unit-by-unit comparison (column 4).

small values (Fig. 7*B*). In contrast, Rayleigh statistics had high values when calculated based on twice the modulation frequency. This indicated that the periodicity in the sFM stimuli was not accurately represented by temporal discharge patterns of this type of response. Despite the different temporal firing patterns produced by sAM and sFM stimuli, the average discharge rate of this unit reached the maximum at the modulation frequency of 16 Hz for both sAM and sFM stimuli (Fig. 4*B*). For the unit shown in Fig. 7*C*, there were no significant stimulus-synchronized discharges for either sAM or sFM stimuli at a modulation frequency of 32 Hz where mean firing rates reached the maximum for both stimuli (Fig. 4*C*). The strongest stimulus-synchronized responses were observed at 16 Hz for sAM and between 8 and 16 Hz for sFM stimuli in this unit (Fig. 7*C*). Additional examples in Fig. 7, *D* and *E*, further demonstrate the lack of significant stimulus-synchronized discharges at modulation frequencies where the units discharged maximally as judged by mean firing rate (Fig. 4, *D* and *E*).

POPULATION PROPERTIES. Figure 8*A* shows distributions of tBMF$_{sAM}$ and tBMF$_{sFM}$ that were analyzed based on calculations at the modulation frequency. The two distributions were statistically indistinguishable (Wilcoxon rank-sum test, $P =$ 0.82). An important property is that the median tBMF of the population is 9.6 Hz for sAM stimuli and 10.0 Hz for sFM stimuli, respectively, which are nearly one octave lower than their counterparts derived from rMTF (median rBMF$_{sAM}$: 22.6

Hz, median rBMF$_{sFM}$: 18.1 Hz, see Table 1). Direct comparison between tBMF$_{sAM}$ and tBMF$_{sFM}$ in the same unit is shown in Fig. 8, *B* and *C*. Similar to the discharge rate-based analysis, a large proportion of recorded units had closely matched tBMFs produced by sAM and sFM stimuli (Fig. 8*B*). The difference between tBMF$_{sAM}$ and tBMF$_{sFM}$ was smaller than the difference between rBMF$_{sAM}$ and rBMF$_{sFM}$ (tBMF: median difference 0.21 octave; rBMF: median difference 0.46 octave; see Table 1). A paired *t*-test showed that there was no significant difference ($P = $ 0.95) between tBMF$_{sAM}$ and tBMF$_{sFM}$ when both were measured in the same units. These data show that both at the level of single neurons and across populations of neurons, there was a large degree of similarity in the preferred modulation frequency as measured by stimulus-synchronized discharges. It should be noted that statistically significant stimulus-synchronized discharges were not detected in a substantial number of units studied (sAM responses: 66/200, 33%; sFM response: 67/142, 47%).

LIMIT ON STIMULUS SYNCHRONIZED DISCHARGES. Another important measure of stimulus-synchronized discharges is the maximum synchronization frequency (f_{max}). This measure indicates the upper limit of stimulus-synchronized discharges in each unit, whereas tBMF defines the modulation frequency at which the strongest discharge synchronization could be induced. Figure 9 shows the distributions of f_{max} for both sAM and sFM stimuli, respectively. The two distributions were
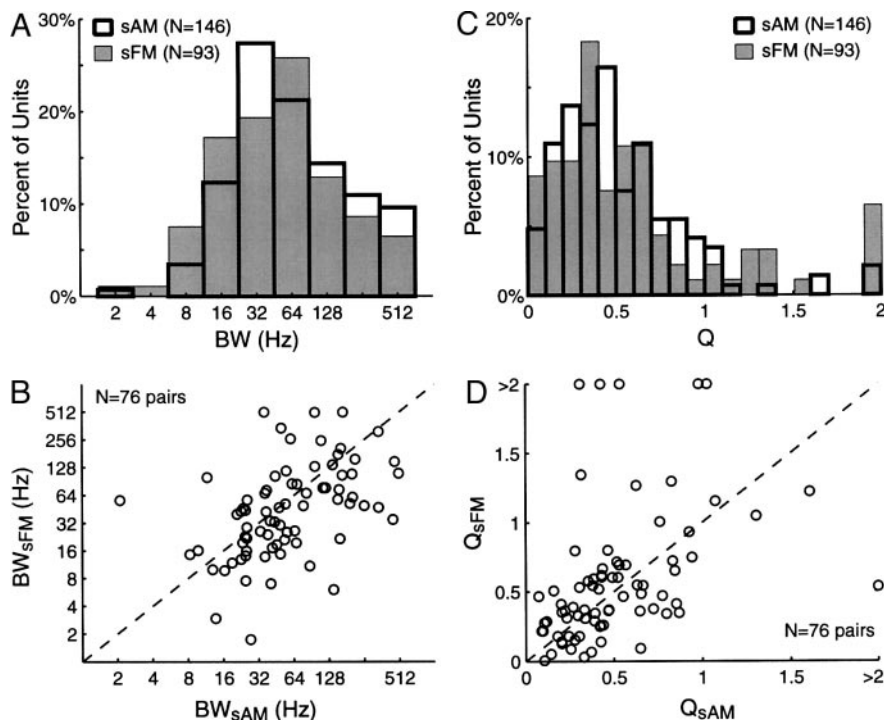
FIG. 6. The tuning width of rMTF. *A*: distributions of half-height BW for sAM (□) and sFM (▨) stimuli, respectively. Two distributions are not statistically different from each other (Wilcoxon rank-sum test, $P = 0.16$). *B*: unit-by-unit comparison between BW_{sAM} and BW_{sFM} (paired *t*-test, $P = 0.27$). *C*: distributions of the sharpness of tuning, $Q = rBMF/BW$, for sAM (□) and sFM (▨) stimuli, respectively (Wilcoxon rank-sum test, $P = 0.77$). *D*: unit-by-unit comparison between Q_{sAM} and Q_{sFM} (paired *t*-test, $P = 0.1$). - - -, slope of 1.0.

statistically indistinguishable (Wilcoxon rank-sum test, $P = 0.97$) and had medians of 34.2 Hz (sAM) and 39.4 Hz (sFM), respectively, which were much higher than their counterparts in tBMF (median $tBMF_{sAM}$: 9.6 Hz, median $tBMF_{sFM}$: 10.0 Hz, see Table 1). Figure 9*B* shows the cumulative distributions of f_{max} for both types of stimuli, which characterizes the upper boundary of stimulus-synchronized activities for the population of recorded units. These curves show how well the A1, as a whole, can represent temporal modulations by temporal discharge patterns. The cumulative distributions of f_{max} for both sAM and sFM stimuli are nearly identical (Fig. 9*B*), indicating the similarity in stimulus-synchronized discharges that resulted from these two classes of stimuli. The curves have low-pass shapes and begin to drop more rapidly above ~16 Hz. The medians of cumulative f_{max} distributions are between 32 Hz and 64 Hz for responses to both types of stimuli (Fig. 9*B*). There were <10% of units that were able to synchronize to modulation waveform at 256 Hz. Further comparison on a unit-by-unit basis between f_{max} measured from sAM and sFM responses is shown in Fig. 9, *C* and *D*. There was a high degree of correlation between f_{max} in individual units as well ($r = 0.7$, Fig. 9*C*). Many units had closely matched f_{max} (Fig. 9*D*). A paired *t*-test showed that there was no significant difference ($P = 0.11$) between f_{max} for sAM and sFM stimuli when both were measured in the same units.

DOUBLING OF SYNCHRONIZATION FREQUENCY. As the example in Figs. 2*B* and 7*B* showed, some units exhibited discharge patterns that were synchronized to twice the modulation frequency ($2f_m$). This was more commonly observed in the responses of sFM stimuli when the frequency component of a stimulus shifted into and out of a unit's excitatory response area during each modulation cycle. In these cases, the synchronization index calculated at $2f_m$ was greater than that calculated at f_m (e.g., Fig. 7*B*). Figure 10 shows the analysis of such cases for both classes of stimuli. There were 36 units (36/93, ~39%

of samples) that exhibited synchronization frequency doubling due to sFM stimuli. In contrast, only a small number of units (7/146, ~5% of samples) were found to show this property with sAM stimuli. In the latter case, the doubling was likely caused by on and off responses to each modulation cycle. For most of the units shown in Fig. 10, f_{max} can be measured using either f_m or $2f_m$. A higher f_{max} value was obtained for most of these units when $2f_m$ was used in the calculation (Fig. 10). In four units, f_{max} could only be measured using $2f_m$ but not f_m in their responses to sAM stimuli, indicating a nearly complete doubling of synchronization frequency (Fig. 10, larger pluses). The doubling of the synchronization frequency did not result in a significant shift of the f_{max} distribution when the entire population of neurons was considered.

Comparison between rate- and synchrony-based modulation frequency selectivity

As illustrated in Figs. 4 and 7, neurons were typically tuned to a higher modulation frequency when measured by average discharge rate than by synchronized discharges. In Fig. 11, we compared rate- and synchrony-based modulation frequency selectivity on a unit-by-unit basis when both rBMF and tBMF could be measured in the same units. For the vast majority of units, rBMF was greater than tBMF, for both sAM (Fig. 11*A*) and sFM (Fig. 11*B*) stimuli. On average, rBMF is more than twice higher than tBMF (Table 2). This means that the average discharge rate reaches the maximum at modulation frequencies as high as where the strongest stimulus-synchronized discharges could be observed. The difference between rBMF and tBMF is statistically significant (sAM: paired *t*-test, $P < 0.001$; sFM: paired *t*-test, $P < 0.001$; Table 2). The correlation coefficient between rBMF and tBMF was small (sAM: 0.31, sFM: 0.09; Table 2). In some units, f_{max} was found at frequencies higher than rBMF as shown in Fig. 11, *C* and *D*. Direct
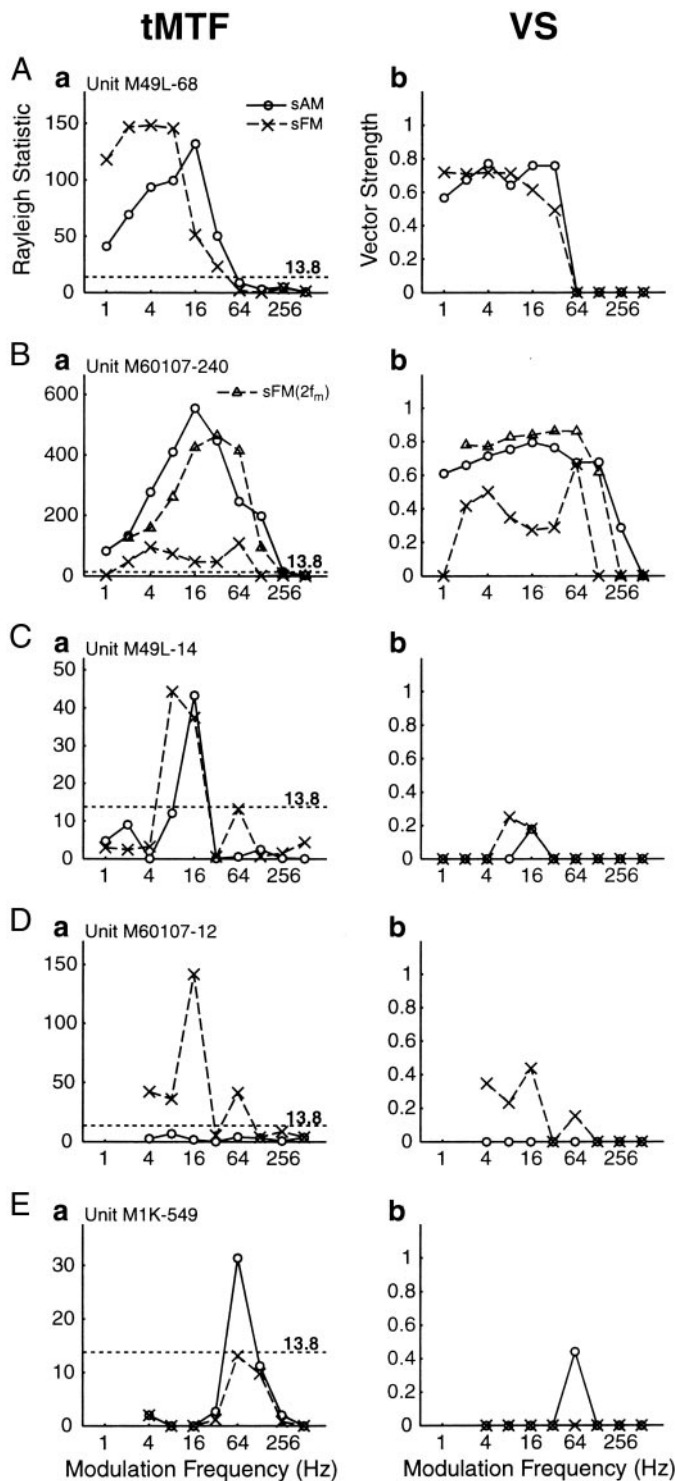
FIG. 7. *A–E, a*: examples of temporal modulation transfer function (tMTF). ··· (equal to 13.8), the threshold for statistically significant stimulus-synchronized responses (Rayleigh test, $P < 0.001$). tMTFs due to sAM (○—○) and sFM (×—×) stimuli are calculated at f_m, except in *B* where an additional tMTF (△ - △) is calculated at the frequency equal to twice the modulation frequency ($2f_m$). The best modulation frequencies (tBMF) calculated from tMTF (see METHODS) for the 5 examples shown are as follows (tBMF$_{sAM}$, tBMF$_{sFM}$). *A*: 6.19, 4.99 Hz. *B*: 13.75, 12.17 Hz. *C*: 16.0, 10.7 Hz. *D*: none, 8.51 Hz. *E*: 64.0, none Hz. *A–E, b*: examples of vector strength (VS) vs. modulation frequency (f_m) profiles for the same group of units shown in *a*. Nonsignificant VS values were set to 0. Units in *A–E* correspond to those shown in Figs. 2 and 3, *A–E*, respectively.
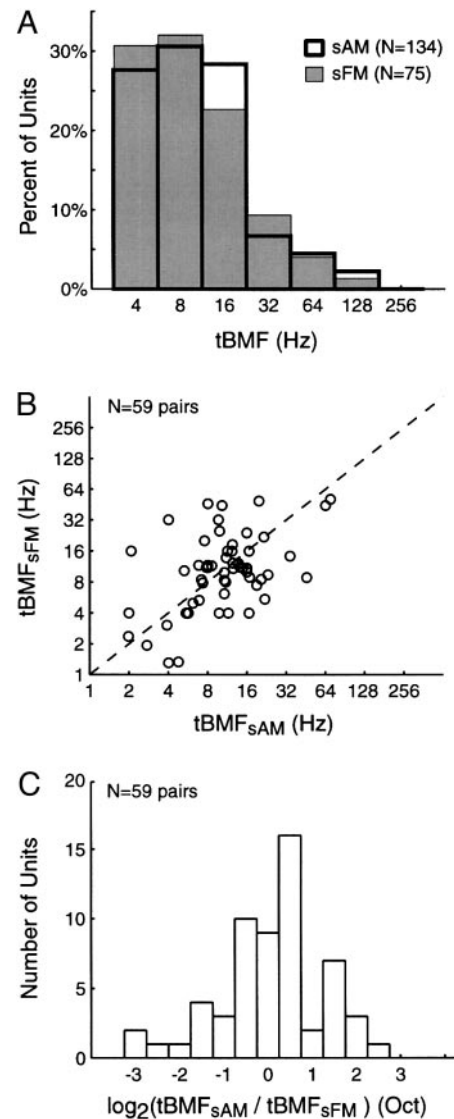
FIG. 8. Population properties for discharge synchrony-based modulation frequency selectivity. *A*: overlapping histograms showing distributions of tBMF derived from sAM (□) and sFM (▨) stimuli, respectively. The binwidths of the histograms are on a base-2 logarithmic scale. The distributions of tBMF$_{sAM}$ and tBMF$_{sFM}$ are not statistically different from each other (Wilcoxon rank-sum test, $P = 0.82$). The means of the 2 distributions are 9.7 Hz (tBMF$_{sAM}$) and 9.2 Hz (tBMF$_{sFM}$) on the base-2 logarithmic scale and 15.6 Hz (tBMF$_{sAM}$) and 14.2 Hz (tBMF$_{sFM}$) on the linear scale, respectively (Table 1, columns 1–3). The medians of the 2 distributions are 9.6 Hz (tBMF$_{sAM}$) and 10.0 Hz (tBMF$_{sFM}$), respectively. *B* and *C*: unit-by-unit comparison between tBMF$_{sFM}$ and tBMF$_{sAM}$ recorded from the same units (Table 1, column 4). This group of 59 units was a subset of the units analyzed in *A*. tBMF$_{sAM}$ and tBMF$_{sFM}$ were not statistically different from each other in the unit-by-unit comparison (paired *t*-test, $P = 0.95$). The correlation coefficient (*r*) between tBMF$_{sAM}$ and tBMF$_{sFM}$ (*B*) was 0.5. The difference between tBMF$_{sAM}$ and tBMF$_{sFM}$ was calculated in octaves as log$_2$(tBMF$_{sAM}$/tBMF$_{sFM}$) (*C*, median = 0.21 octaves). - - -, slope of 1.0.

comparisons showed that f_{max} did not differ significantly from rBMF for sAM responses (paired *t*-test, $P = 0.03$, Table 2), although a significant difference was found for sFM responses (paired *t*-test, $P < 0.01$, Table 2). Correlation between f_{max} and rBMF was poor (sAM: 0.20, sFM: 0.01; Table 2). These comparisons showed that rBMF, a discharge rate-based measurement, of a unit was not significantly correlated with discharge synchrony-based measurements (tBMF and f_{max}). In
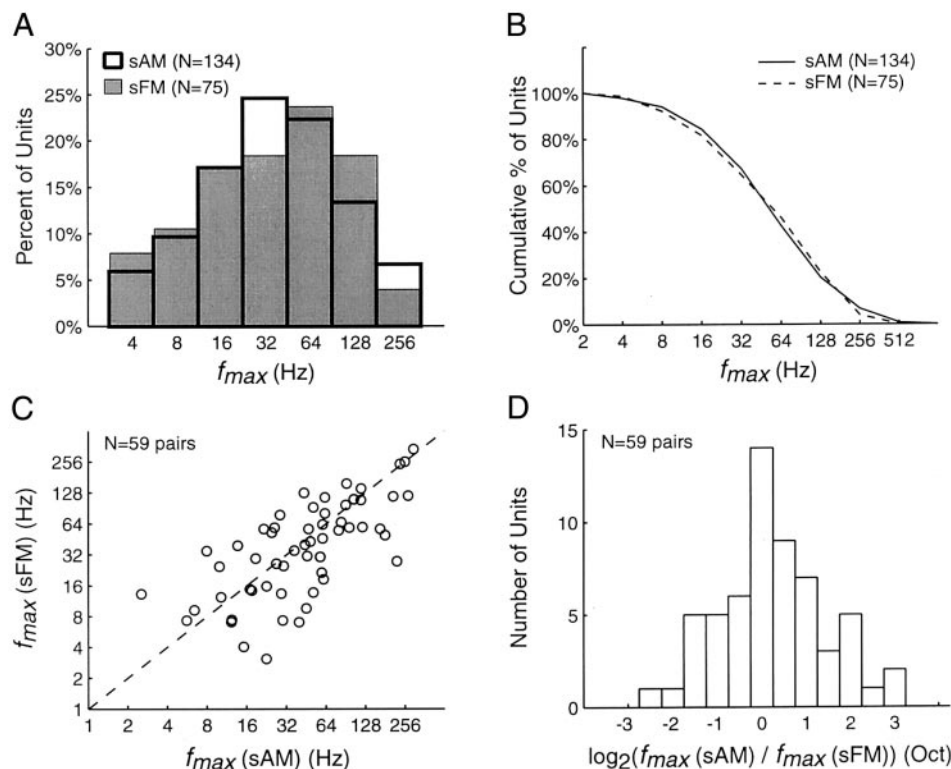
FIG. 9. Population properties for maximum synchronization frequency (f_{max}). *A*: overlapping histograms showing distributions of f_{max} derived from sAM ($\square$) and sFM ($\boxdot$) stimuli, respectively. The binwidths of the histograms are on a base-2 logarithmic scale. The distributions of f_{max}(sAM) and f_{max}(sFM) are not statistically different from each other (Wilcoxon rank-sum test, $P = 0.97$). The means of the 2 distributions are 34.2 Hz (sAM) and 32.9 Hz (sFM) on the base-2 logarithmic scale and 58.9 Hz (sAM) and 57.4 Hz (sFM) on the linear scale, respectively (Table 1, columns 1–3). The medians of the 2 distributions are 34.2 Hz (sAM) and 39.4 Hz (sFM), respectively. *B*: cumulative distributions of f_{max} based on the same populations of units shown in *A*. *C* and *D*: unit-by-unit comparison between f_{max}(sAM) and f_{max}(sFM) recorded from the same units (Table 1, column 4). This group of 59 units was a subset of the units analyzed in *A*. f_{max}(sAM) and f_{max}(sFM) were not statistically different from each other in the unit-by-unit comparison (paired *t*-test, $P = 0.11$). The correlation coefficient (*r*) between f_{max}(sAM) and f_{max}(sFM) (*C*) was 0.7. The difference between f_{max}(sAM) and f_{max}(sFM) was calculated in octaves as $\log_2(f_{max}(\text{sAM})/f_{max}(\text{sFM}))$ (*D*, median = 0.14 octaves). - - -, slope of 1.0.

contrast, tBMF and f_{max} were highly correlated while they differed significantly (sAM: paired *t*-test, $P < 0.001$, $r = 0.65$; sFM: paired *t*-test, $P < 0.001$, $r = 0.70$; Table 2). f_{max} is more than three times higher than tBMF (Fig. 11, *E* and *F,* Table 2). The distributions in Fig. 11 again showed the similarity between responses to sAM and sFM stimuli.

Comparisons between neural populations

In Fig. 12, rate- and synchrony-based response measures were compared between different populations of units. The recorded units were partitioned into BP and non-BP groups in our analyses (see METHODS). Averaged discharge rates are plotted versus modulation frequency for both of these groups as well as for all units in Fig. 12, *A* (sAM) and *B* (sFM). The population-averaged discharge rate profiles of the BP-group showed a maximum between 16 and 32 Hz of modulation frequency (Fig. 12, *A*, and *B*, $\square$—$\square$), similar to that observed in the distributions of rBMF measured from individual units (Fig. 5*A*). This feature can also be seen when the responses of the entire population are averaged (Fig. 12, *A* and *B*, - - -). These observations indicate that not only were there more units tuned to modulation frequencies in the range of 16–32 Hz, the A1 responded collectively more strongly to this range of modulation frequencies than to lower or higher modulation frequencies. The profiles of the non-BP group, however, were flat between 4 and 64 Hz and showed an increase in discharge rate at higher modulation frequencies (Fig. 12, *A* and *B*, $\triangle$ - $\triangle$).

Figure 12, *C* and *D,* showed the proportion of units that exhibited statistically significant Rayleigh statistic at each modulation frequency for the three groups of units. The highest percentages of units with significant synchronized discharges were between 4- and 16-Hz modulation frequency and were centered near 8 Hz (sAM: 48%, sFM: 32%), consistent with the

distribution of tBMF of individual units (Fig. 8*A*). Less than half of all sampled units showed stimulus-synchronized discharges at any tested modulation frequency. These profiles reflect the overall strength of stimulus-synchronized discharges across modulation frequency that are evoked by sAM and sFM stimuli. The low percentages of units with synchronized discharges in the non-BP group were partially explained by the relatively low response magnitudes of these units (Fig. 12, *A* and *B*). Data in Fig. 12 demonstrate that the units belonging to the BP group carry far more information than units of the
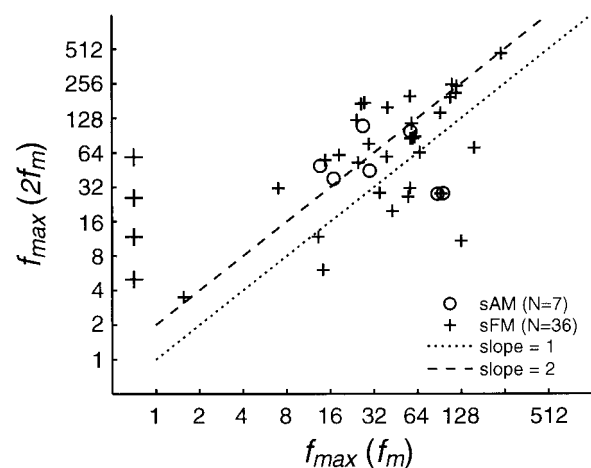

FIG. 10. Unit-by-unit comparison between $f_{max}(f_m)$ and $f_{max}(2f_m)$ calculated from tMTF based on f_m and $2f_m$, respectively, for a subpopulation of units that had doubling of synchronization frequency (see METHODS). f_{max} calculated from sAM and sFM responses are indicated by circles and pluses, respectively. Diagonal dotted line, slope of 1; dashed line, slope of 2. Units with significantly synchronized discharges at $2f_m$ (as judged by Rayleigh statistic calculated at $2f_m$) but not at f_m are shown as larger pluses alongside the ordinate.
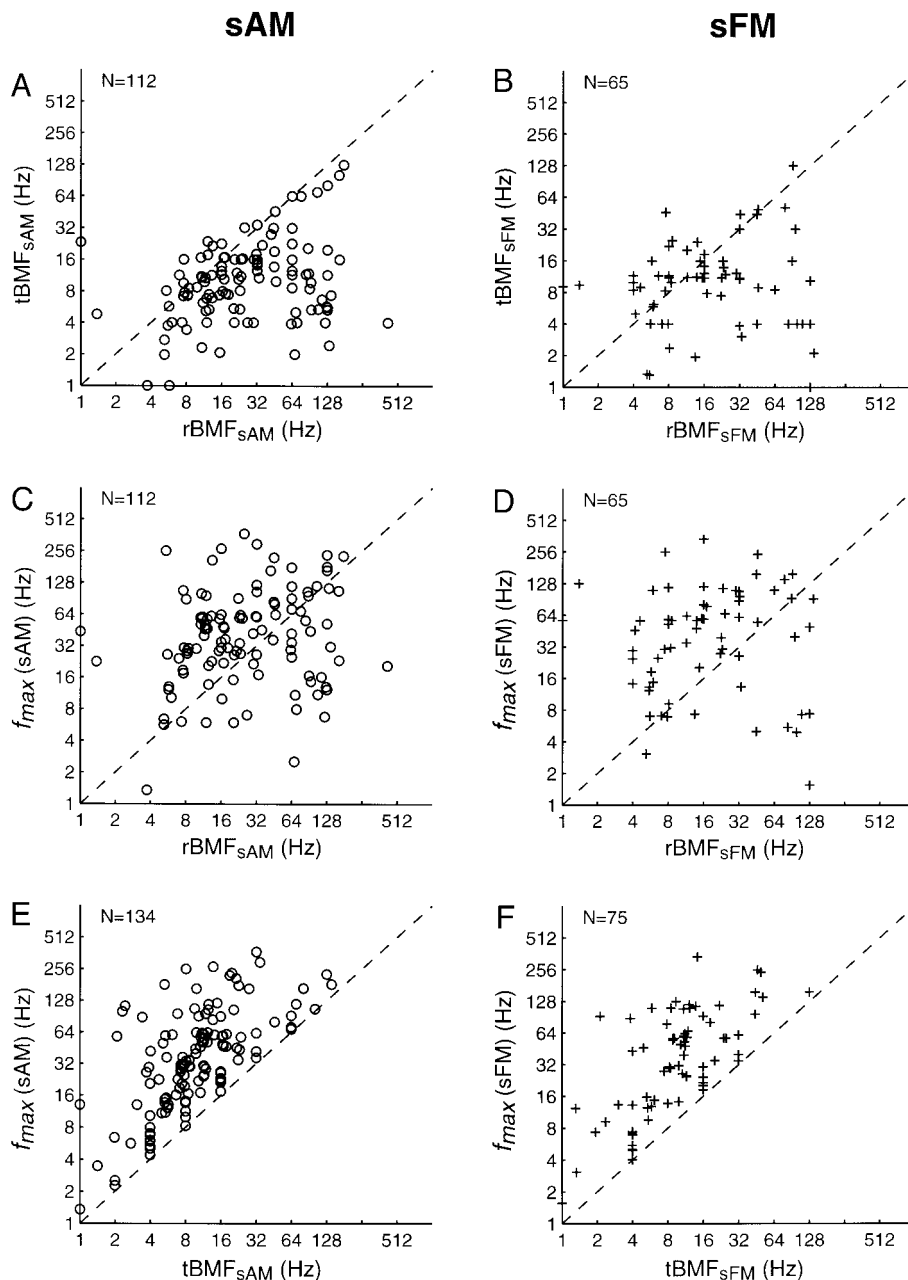
FIG. 11. Comparison between rate- and synchrony-based modulation selectivity evaluated on the basis of individual units. A and B: tBMF is plotted vs. rBMF for sAM stimuli (A: $n = 112$) and sFM stimuli (B: $n = 65$). tBMF differed significantly from rBMF for both sAM and sFM stimuli (paired t-test, $P < 0.001$; Table 2). C and D: comparison between f_{max} and rBMF for sAM stimuli (C: $n = 112$, paired t-test, $P = 0.03$) and sFM stimuli (D: $n = 65$, paired t-test, $P < 0.01$), respectively. E and F: comparison between f_{max} and tBMF for sAM stimuli (E: $n = 134$) and sFM stimuli (F: $n = 75$). f_{max} differed significantly from tBMF for both sAM and sFM stimuli (paired t-test, $P < 0.001$; Table 2). - - -, slope of 1 in all plots.

non-BP group in terms of both rate- and synchrony based representations of modulation frequency.

Dependency of modulation-frequency selectivity on stimulus parameters

DEPENDENCE OF MTF ON SOUND LEVEL. Rate-level functions were nonmonotonic for narrowband stimuli (e.g., pure and modulated tones) in most units that we studied in awake marmosets (Wang et al. 1999). We tested in a subset of units the dependence of modulation selectivity on sound level and found that the shape of rMTF and, to a lesser extent, the shape of tBMF was relatively invariant across supra-threshold sound levels. Figure 13, A and B, shows examples of sAM responses from two units. In each case, while changing sound level resulted in changes in firing rates, both rBMF and tBMF remained largely unchanged across a wide range of sound

levels (Fig. 13, A and B). Notice that the rate-level relationship was nonmonotonic for the units shown in Fig. 13, A and B. Figure 13C shows an analysis of sound-level dependence of rBMF (Fig. 13Ca) and tBMF (Fig. 13Cb) over a population of units. For most of the units tested, changing sound level did not result in large changes in rBMF or tBMF. Similar observations were obtained with sFM responses as shown in Fig. 14, in individual examples (Fig. 14, A and B) as well as in a population analysis (Fig.14C). The unit shown in Fig. 14B had large changes in tMTF but much smaller changes in rMTF across sound levels. In general, the shape of rMTF, and consequently rBMF, tended to be more resistant to changes in sound level than did tMTF and tBMF. These data showed that although changing sound level may change the peak firing rate and the sharpness of tuning to the preferred modulation frequency, it usually did not result in significant shift of MTF. There ap-

TABLE 2. *Unit-by-unit comparison between response measures*

	rBMF vs. tBMF		rBMF vs. f_{max}		tBMF vs. f_{max}	
	rBMF, Hz	tBMF, Hz	rBMF, Hz	f_{max}, Hz	tBMF, Hz	f_{max}, Hz
sAM	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):
	25.6 [8.4, 78.2]	9.9 [4.1, 24.0]	25.6 [8.4, 78.2]	38.3 [13.4, 109.9]	9.7 [3.9, 24.1]	34.2 [11.3, 103.1]
	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):
	45.8 ± 56.1	15.2 ± 19.2	45.8 ± 56.1	63.1 ± 68.4	15.6 ± 21.4	58.9 ± 66.2
	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:
	22.9 [11.7, 64.0]	10.1 [5.4, 16.0]	22.9 [11.7, 64.0]	40.9 [21.1, 75.8]	9.6 [5.4, 16.0]	34.2 [17.0, 63.6]
	$P < 0.001$, $r = 0.31$, $n = 112$		$P = 0.03$, $r = 0.20$, $n = 112$		$P < 0.001$, $r = 0.65$, $n = 134$	
sFM	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):	$\mu \pm$ std (log):
	17.1 [5.4, 54.5]	9.4 [3.7, 24.3]	17.1 [5.4, 54.5]	36.8 [11.4, 119.1]	9.2 [3.7, 23.0]	32.9 [10.4, 104.4]
	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):	$\mu \pm$ std (linear):
	32.0 ± 37.0	14.8 ± 18.6	32.0 ± 37.0	63.3 ± 64.2	14.2 ± 17.6	57.4 ± 61.7
	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:	median [25%, 75%]:
	16 [7.2, 33.4]	10.7 [4.0, 16.0]	16.0 [7.2, 33.4]	49.7 [14.5, 92.2]	10.0 [4.0, 16.0]	39.4 [13.4, 79.4]
	$P < 0.001$, $r = 0.09$, $n = 65$		$P < 0.01$, $r = 0.01$, $n = 65$		$P < 0.001$, $r = 0.70$, $n = 75$	

Statistics shown are based on the data presented in Fig. 11. $\mu \pm$ std (log) is as defined in Table 1. P is the paired t-test score. r is the correlation coefficient based on the log2-scale. n is number of units.

peared to be, however, greater changes due to sound level in sFM responses than in sAM responses.

DEPENDENCE OF MTF ON MODULATION DEPTH. Modulation depth was another important parameter that affected the responsiveness of a unit to modulated sounds. In general, increasing modulation depth of a sAM stimulus at or near rBMF always led to a monotonic change in discharge rate (increased or saturated). For the unit shown in Fig. 15A, the maximum firing rate of the rMTF increased from ~13 to ~38 spikes/s as modulation depth of the sAM stimuli was increased from 50 to 100%. However, the rBMF remained unchanged near 32 Hz. This was also true for the synchronization measure, with tBMF near 16 Hz (Fig. 15Ab). Similar properties can be seen in

another example in Fig. 15B. Figure 15C shows a population analysis that further strengthens these observations regarding responses to sAM stimuli.

A unit's responses to changes of the modulation depth of sFM stimuli were, however, more complicated. Unlike sAM stimuli, where changing modulation depth only altered the magnitude but not spectral spread of the side bands, changing sFM modulation depth may result in moving the side bands into or out of the excitatory and inhibitory regions of the response area of a unit. As a result, there was generally a more complex relationship between the shape of rMTF and the modulation depth of sFM stimuli. Despite these factors, the shape of MTFs and consequently rBMF remained largely un-
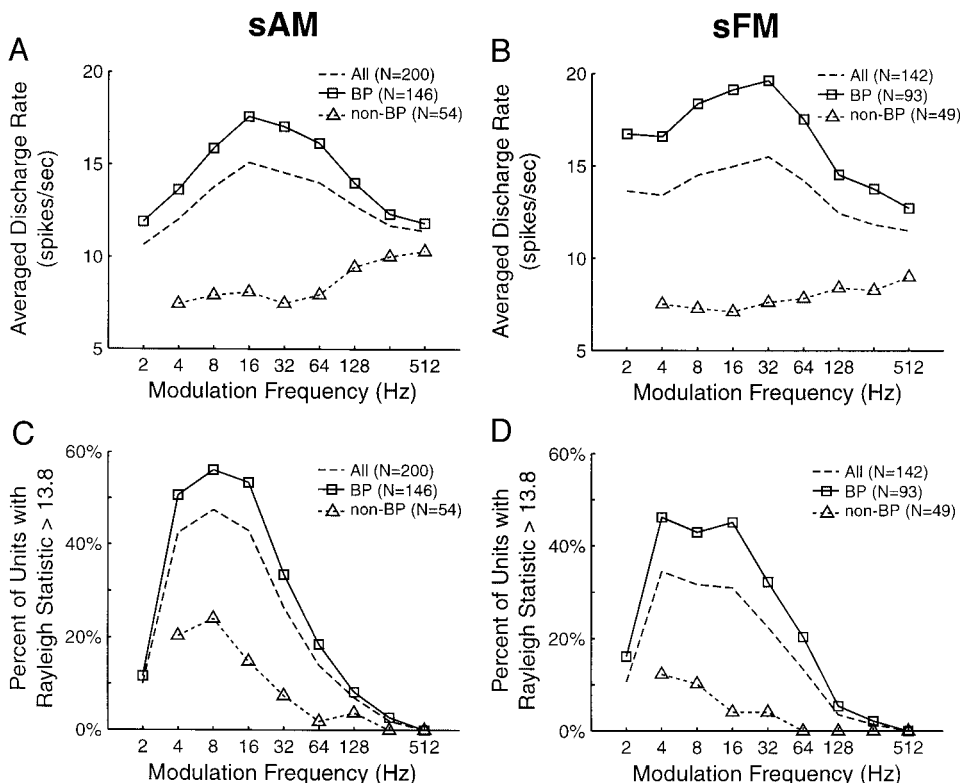

FIG. 12. Comparisons of rate- and synchrony-based response measures between populations of units. *A* and *B*: discharge rates averaged over 3 groups of units, all (- - -), band-pass (BP, □—□) and non-BP (△—△) group, respectively, are plotted vs. modulation frequency for sAM (*A*) and sFM (*B*) stimuli. Spontaneous rates were not subtracted in the calculations. *C* and *D*: percent of units with statistically significant Rayleigh statistic (>13.8) is plotted vs. modulation frequency. Three groups of units, all (- - -), BP (□—□), and non-BP (△—△) group, respectively, are shown for sAM (*C*) and sFM (*D*) stimuli.
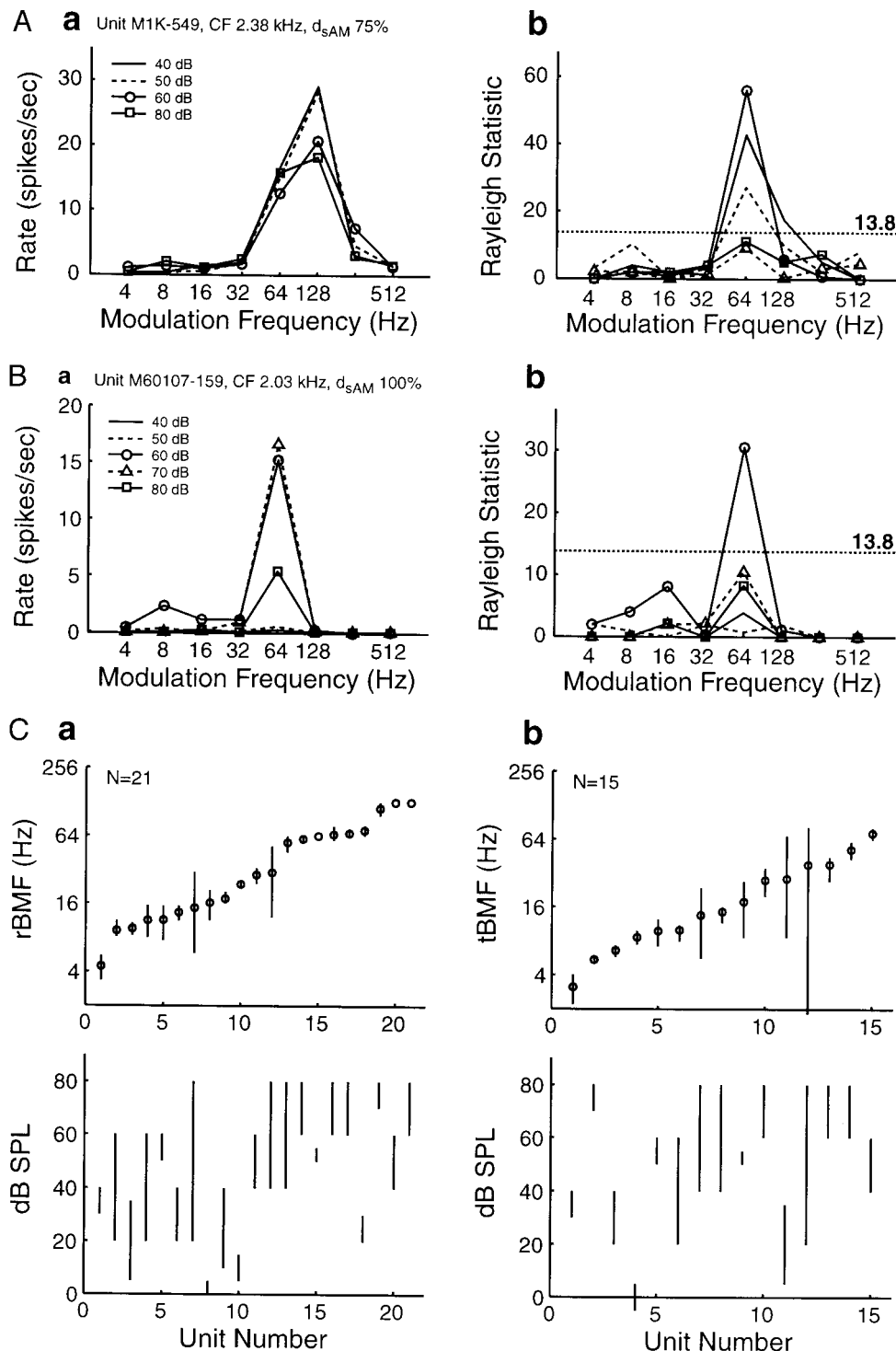
sAM

FIG. 13. Effects of sound level on rMTF (*left, a*) and tMTF (*right, b*) for sAM stimuli. *A* and *B*: examples of MTFs measured at multiple sound levels in 2 representative units. $\cdots$ (equal to 13.8), the statistical significance threshold for Rayleigh test ($P < 0.001$). Stimulus parameters are indicated on each plot. *C*: summary of population properties to changes in sound level for rBMF (*a*) and tBMF (*b*), respectively. *Top*: the mean ($\circ$) and the range of minimum and maximum ($|$) rBMF (*a*) or tBMF (*b*) measured in each unit. Units are ordered by their mean rBMF or tBMF, respectively (from low to high). Corresponding ranges of sound levels tested for each unit are shown in *bottom* plots. d_{sAM}, modulation depth for sAM stimuli. At least two sound levels were tested for each data point shown in *C*.

changed when the sFM modulation depth was varied, as can be seen by an individual example in Fig. 16*Aa* and a population analysis in Fig. 16*Ca*. Similar properties were observed for tMTF and tBMF (Fig. 16, *Ab* and *Cb*). The example in Fig. 16*Ba* showed that the peak of rMTF was shifted toward lower modulation frequency as modulation depth was increased from 1,024 to 2,048 and then to 4,096 Hz. This shift could be due to the fact that side band inhibitions may be evoked at these higher modulation depths and may explain lower rBMF$_{sFM}$ than rBMF$_{sAM}$ in units with high rBMF$_{sAM}$ values as shown in Fig. 5*B* (also see Table 1). The corresponding shifts in tMTF were smaller in this example (Fig. 16*Bb*).
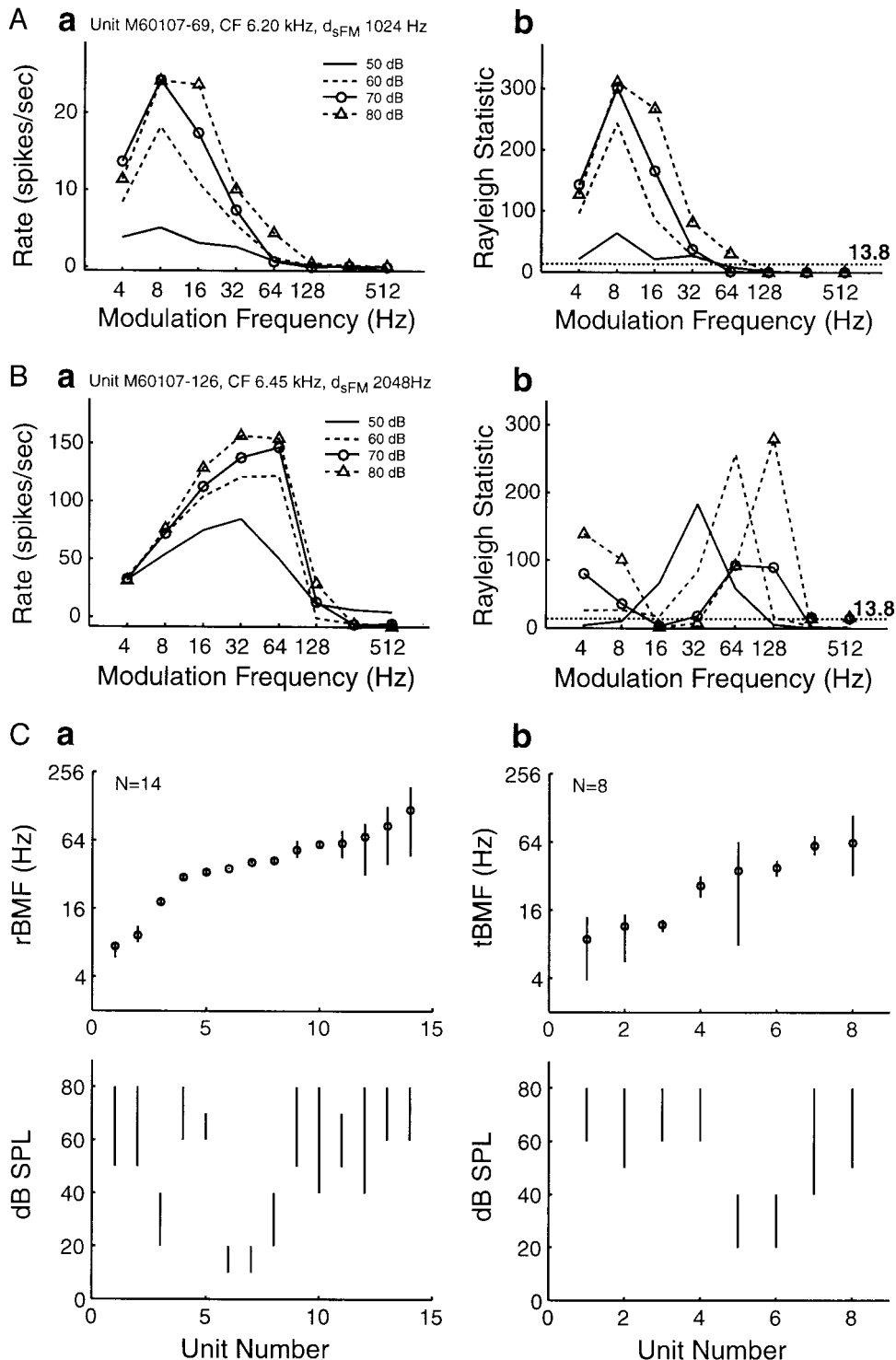
sFM

FIG. 14. Effects of sound level on rMTF and tMTF for sFM stimuli. The format is the same as Fig. 13. d_{sFM}, modulation depth for sFM stimuli.

Temporal modulation is essential to drive many cortical neurons

Temporal modulation (in amplitude or frequency) was essential to drive many units that were unresponsive or only weakly responsive to pure tones. Figure 17, *A* and *B,* shows two representative examples of a class of units that required sufficient AM to fire. The unit in Fig. 17*A* gave only onset discharges at zero modulation depth (equivalent to a CF tone) but gave sustained discharges when the modulation depth of the sAM stimulus was raised to $\geq 50\%$. Figure 17*B* shows another example that had offset discharges at modulation depths $<70–80\%$ but fired continuously throughout the stimulus duration at greater modulation depths. Discharge rate versus modulation depth functions from a group of such units
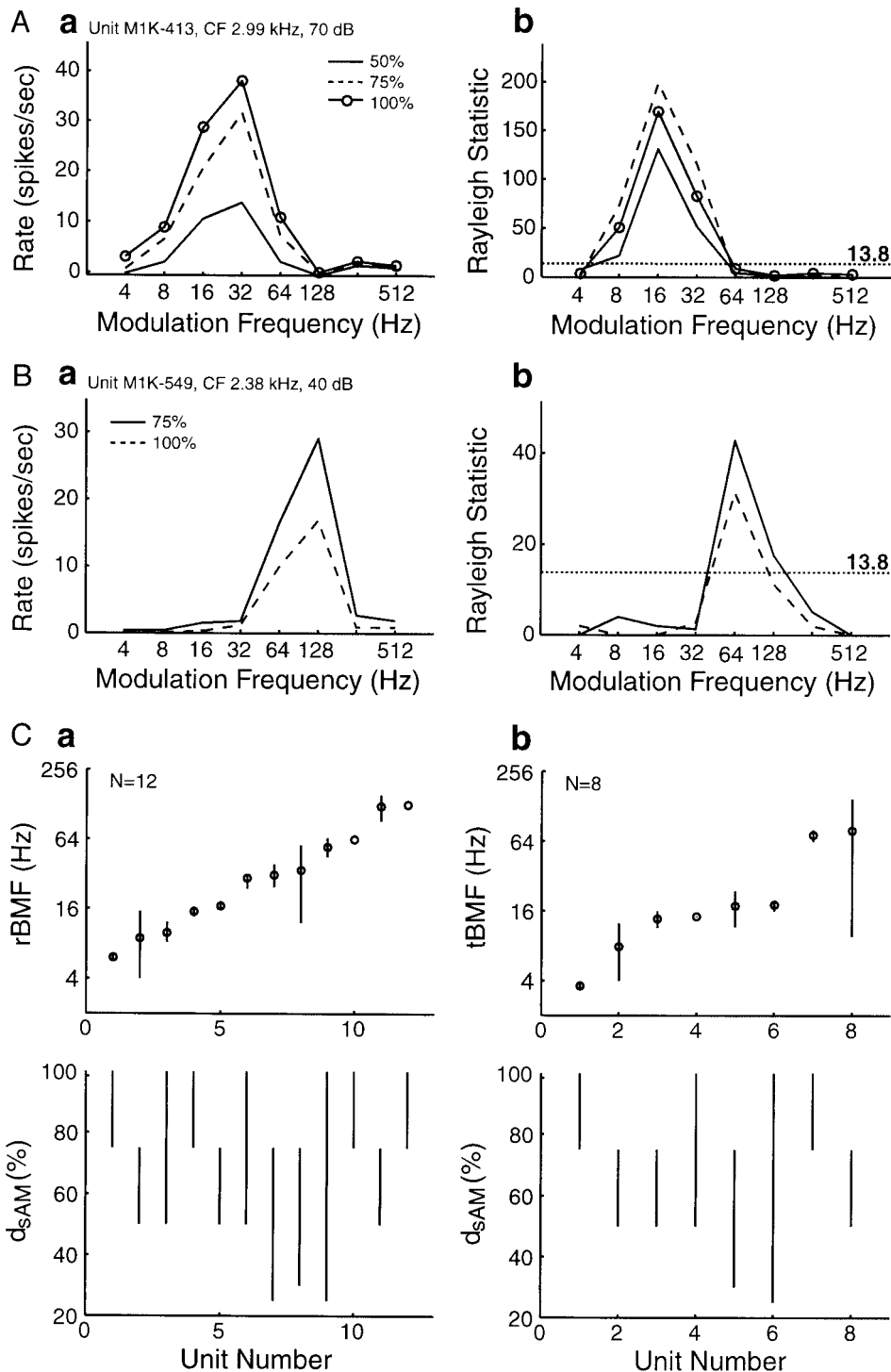
sAM

FIG. 15. Effects of modulation depth on rMTF and tMTF for sAM stimuli. The format is the same as Fig. 13. d_{sAM}, modulation depth for sAM stimuli.

is shown in Fig. 17*C*. These units had weak or no responses to sAM stimuli at zero modulation depth, i.e., unmodulated CF tones. Some units did not respond to pure tones until they were sufficiently modulated in frequency, as shown by examples in Fig. 17, *D–F*. In contrast to sAM stimuli, sFM stimuli generally produced maximum firing rate at a particular modulation depth in each unit. Increasing modulation depth beyond this optimal depth could result in a reduction in firing rate (Fig.

17*F*), presumably due to the recruitment of flanking inhibitions by expanded spectral side bands.

Evidence of inhibition in shaping modulation frequency selectivity

Inhibition appeared to play a role in limiting the range of modulation frequency that a unit would respond and thus
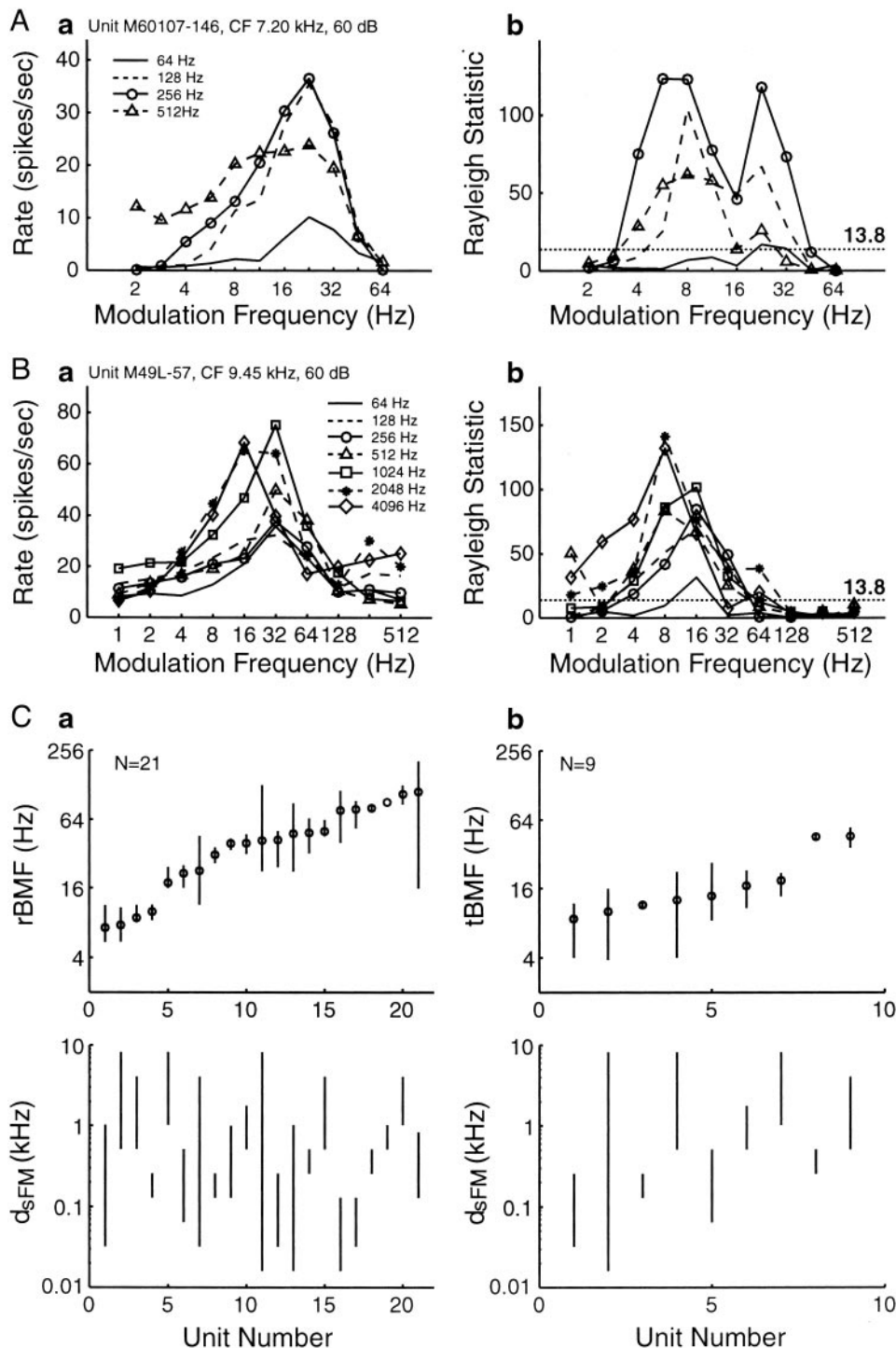
sFM

FIG. 16. Effects of modulation depth on rMTF and tMTF for sFM stimuli. The format is the same as Fig. 13. d_{sFM}, modulation depth for sFM stimuli.

contributed to the observed modulation frequency selectivity. In the examples shown in Fig. 2, *C* and *D,* the responses of the units were clearly inhibited at some modulation frequencies higher than rBMF for both sAM and sFM stimuli. In some cases, inhibitory effects were observed at modulation frequencies below rBMF (e.g., Figs. 2*Da* and 4*D*). In general, inhibition was usually observed at both extremes of a rMTF and was often stronger at the high-frequency side. However, inhibition

could only be detected in units that had spontaneous discharges. For the units that had nearly zero spontaneous rate (e.g., Fig. 2*B*) one cannot be sure that the disappearance of discharges after onset responses at a modulation frequency of 512 Hz was due to inhibition or not. In Fig. 18*A*, we analyzed the percentage of units that showed significant inhibition below spontaneous discharge rate in their sAM responses at each tested modulation frequency. The percentage of units with
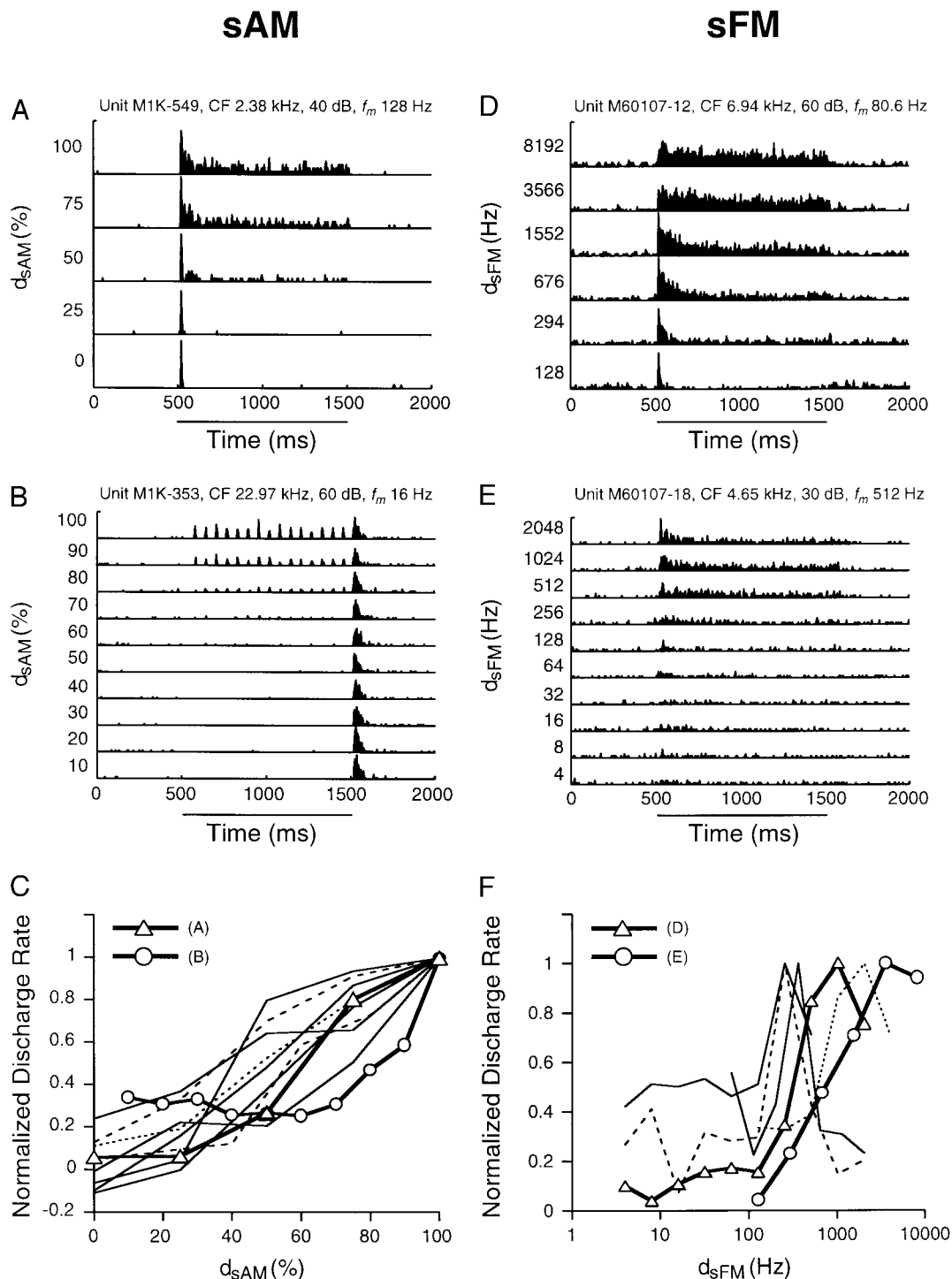
FIG. 17. Dependence of cortical responses on amplitude or frequency modulation. Examples of units that responded weakly to unmodulated tones but strongly to sAM (*left, A–C*) or sFM (*right, D–F*) stimuli with sufficient modulation depths. *A* and *B*: 2 representative examples of units that responded to sAM stimuli during the stimulus duration only at sufficiently large modulation depths. Onset (*A*) or offset (*B*) responses were present at all modulation depths. *C*: normalized discharge rate is plotted vs. modulation depth for a group of units. Mean firing rate at 100% modulation depth was used for the normalization. Maximum firing rate in each unit was usually achieved at or near the largest modulation depth tested. *D* and *E*: 2 representative examples of units that responded to sFM stimuli during the stimulus duration only at sufficiently large modulation depths. *F*: firing rate, normalized by the maximum discharge rate for each unit, is plotted vs. modulation depth for a group of units. In *A, B, D,* and *E,* the frequency of the carrier tone was set at a unit's CF. The modulation frequency was either at or near rBMF of each unit. f_m, modulation frequency.

inhibition was smallest (~2%) near 32-Hz modulation frequency, which was consistent with the observation that there were more units with rBMF near this modulation frequency

than anywhere else (Fig. 5*A*). The largest percentage of units with inhibition was at the highest modulation frequency tested, 512 Hz (~13%, Fig. 18*A*). There were fewer units with inhi-
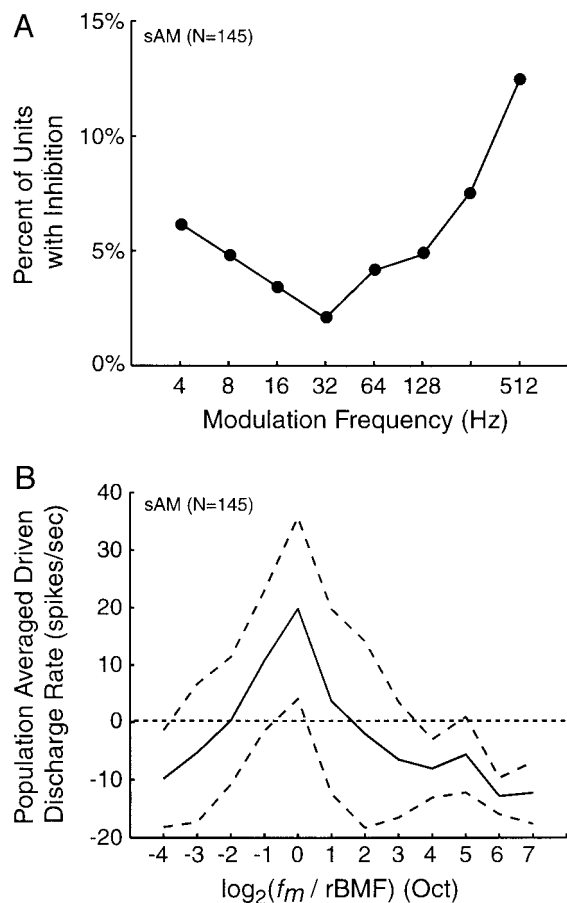
FIG. 18. Evidence of inhibition in shaping modulation-frequency selectivity for sAM stimuli. 145 units with discharge rates that were significantly less than spontaneous rates ($P < 0.05$, Wilcoxon rank-sum test), at one or more modulation frequency, were analyzed as a population. *A*: proportion of units that displayed inhibition in their responses at each tested modulation frequency. *B*: population-averaged rMTF with modulation frequency normalized by the rBMF of each unit (—, mean; - - -, 1 SD from the mean), plotted on a base-2 logarithmic scale (octave). Negative firing rates are relative to spontaneous activity and indicate the presence of inhibition.

bition at the lowest modulation frequency tested, 4 Hz (~6%). Figure 18*B* shows the population-averaged mean firing rate, in response to sAM stimuli, as a function of modulation frequency, based on rMTF from individual units that were aligned by their rBMF. At the highest and lowest modulation frequencies tested, both the mean rate and 1 SD above the mean rate fell below the spontaneous activity level (Fig. 18*B*). This indicated that there was a stronger presence of inhibition at these modulation frequencies. The shapes of the distributions shown in Fig. 18 are reminiscent of the shape of rBMF distributions in Fig. 5*A*, suggesting that inhibition may play a role in shaping overall modulation frequency selectivity observed in A1.

Relationship between modulation-frequency selectivity and response latency and CF

In Fig. 19, all three response measures are plotted, respectively, versus minimum response latency for both sAM and sFM stimuli. The minimum response latency was calculated from a unit's responses to modulated tones across all modulation frequencies tested (see METHODS). There was no clear

correlation between rBMF and the latency for both sAM and sFM stimuli (Fig. 19*A*). rBMF did not differ significantly between units with shorter latencies and those with longer latencies (Wilcoxon rank-sum test, $P = 0.03$, Table 3). However, there appeared to be a weak inverse relationship between tBMF and the latency (Fig. 19*B*). Units with higher tBMF
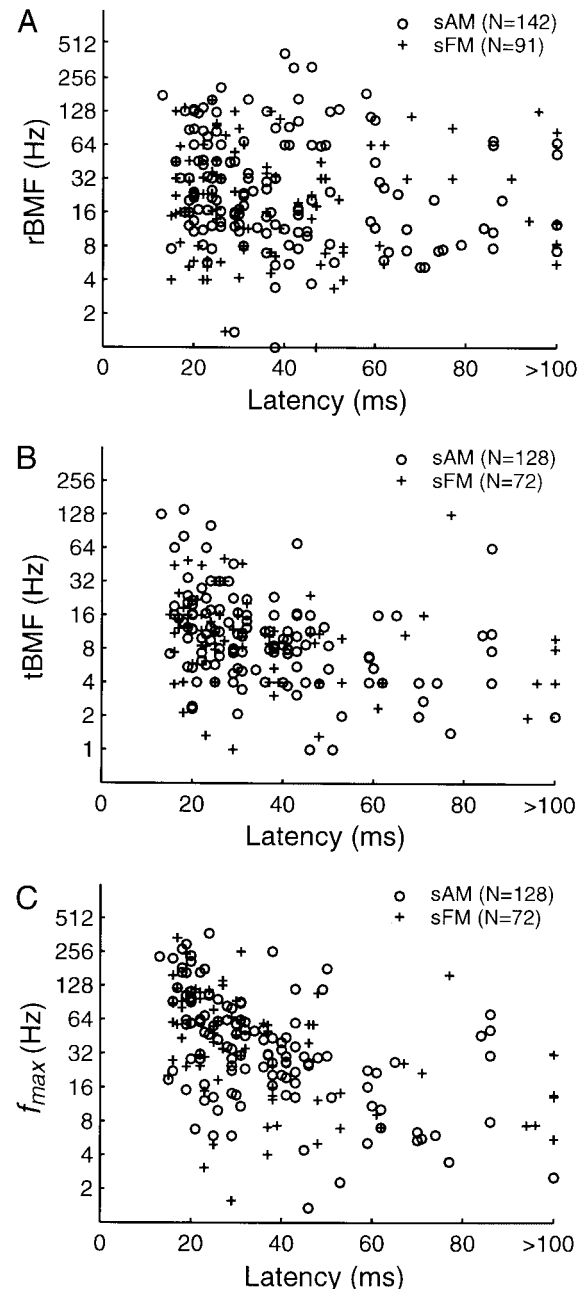

FIG. 19. Relationships between measures of MTF and the minimum response latency are shown for both sAM (○) and sFM (+) data. *A*: rBMF, *B*: tBMF, and *C*: f_{max}. The minimum response latency is the minimum latency of responses at all tested modulation frequencies (see METHODS) and is calculated for sAM and sFM responses separately. Statistical comparisons were made between 2 groups of units separated by the mean latency (see Table 3). Significant differences were found between the medians calculated from units with latency less than the median latency and those with latency greater than the median latency for f_{max} (sAM: $P < 0.001$, sFM: $P < 0.01$, Wilcoxon rank-sum test) and tBMF (sAM: $P < 0.001$, sFM: $P = 0.018$, Wilcoxon rank-sum test) but not for rBMF (sAM: $P = 0.03$, sFM: $P = 0.34$, Wilcoxon rank-sum test). See Table 3.

TABLE 3. *Dependency on latency and characteristic frequency (CF)*

	Latency		CF	
	sAM	sFM	sAM	sFM
rBMF, Hz	Latency $< m_{Lat}$: 24.6 ($n = 71$) Latency $\geq m_{Lat}$: 16.0 ($n = 71$) (m_{Lat}: 33.5 ms) $P = 0.03$	Latency $< m_{Lat}$: 23.2 ($n = 45$) Latency $\geq m_{Lat}$: 15.0 ($n = 46$) (m_{Lat}: 31.0 ms) $P = 0.34$	CF $< m_{CF}$: 32.0 ($n = 73$) CF $\geq m_{CF}$: 17.0 ($n = 73$) (m_{CF}: 6.82 kHz) $P = 0.03$	CF $< m_{CF}$: 28.2 ($n = 46$) CF $\geq m_{CF}$: 16.2 ($n = 47$) (m_{CF}: 6.91 kHz) $P = 0.2$
tBMF, Hz	Latency $< m_{Lat}$: 12.5 ($n = 62$) Latency $\geq m_{Lat}$: 7.9 ($n = 66$) (m_{Lat}: 31.0 ms) $P < 0.001$	Latency $< m_{Lat}$: 12.0 ($n = 36$) Latency $\geq m_{Lat}$: 8.6 ($n = 36$) (m_{Lat}: 29.5 ms) $P = 0.018$	CF $< m_{CF}$: 11.5 ($n = 67$) CF $\geq m_{CF}$: 8.0 ($n = 67$) (m_{CF}: 6.95 kHz) $P = 0.015$	CF $< m_{CF}$: 11.0 ($n = 37$) CF $\geq m_{CF}$: 8.5 ($n = 38$) (m_{CF}: 6.99 kHz) $P = 0.14$
f_{max}, Hz	Latency $< m_{Lat}$: 60.7 ($n = 62$) Latency $\geq m_{Lat}$: 26.7 ($n = 66$) (m_{Lat}: 31.0 ms) $P < 0.001$	Latency $< m_{Lat}$: 60.7 ($n = 36$) Latency $\geq m_{Lat}$: 23.3 ($n = 36$) (m_{Lat}: 29.5 ms) $P < 0.01$	CF $< m_{CF}$: 30.8 ($n = 67$) CF $\geq m_{CF}$: 36.8 ($n = 67$) (m_{CF}: 6.95 kHz) $P = 0.8$	CF $< m_{CF}$: 34.9 ($n = 37$) CF $\geq m_{CF}$: 41.2 ($n = 38$) (m_{CF}: 6.99 kHz) $P = 0.86$

Statistics shown are based on the data presented in Fig. 19 (latency) and Fig. 20 (CF). Values given are medians of corresponding groups of units. P is the Wilcoxon rank sum test score. m_{Lat} and m_{CF} are median latency and CF, respectively.

tended to have shorter latency, and those with the longest latency ($\sim$100 ms) mostly had the lowest tBMF (Fig. 19*B*). The median tBMF of units with latencies shorter than the median latency was greater than that of the units with latencies longer than the median latency (sAM: 12.5 vs. 7.9 ms, $P < 0.001$; sFM: 12.0 vs. 8.6 ms, $P = 0.018$; see Table 3). f_{max} showed the strongest dependency on the latency (Fig. 19*C*, sAM: 60.7 vs. 26.7 ms, $P < 0.001$; sFM: 60.7 vs. 23.3 ms, $P < 0.01$; see Table 3). The inverse relationship between f_{max} and the latency was clearer than that observed with tBMF. In contrast, there appeared to be no dependency between the three response measures and the CF (Fig. 20). The medians calculated from units with CF less than the median CF were not statistically different from those from units with CF greater than the median CF (Table 3), suggesting that the observed modulation selectivity is a likely fundamental property of the A1 across the tonotopic axis.

DISCUSSION

Common temporal processing mechanism

In the reported experiments, temporal modulations were introduced in the amplitude or frequency domain. While these two classes of stimuli differed significantly in their spectral contents, they shared the same modulation waveform: a sinusoid with a particular temporal modulation frequency that gave rise to the perceived time-varying property of these sounds. Results presented in this report clearly demonstrated that there was a high degree of similarity between cortical responses to sAM and sFM stimuli. Auditory cortical neurons could be selective to particular modulation frequencies assessed by their mean firing rate or by discharge synchrony. The selectivity was similar regardless of whether the temporal modulation was created in the amplitude or frequency domain, suggesting that it was the temporal modulation frequency per se, and not the spectral contents, to which the neurons were selective. Because amplitude and frequency modulations are produced along different stimulus dimensions, the match between major aspects of sAM and sFM responses (e.g., rBMF, BW, tBMF, and f_{max}) suggests a *shared* representation of temporal modulations by cortical neurons. Our findings support the notion that there exists a common mechanism in cortical neurons for extracting

temporal profiles from a variety of complex sounds with different spectral contents (Wang et al. 2001). In fact, our recent studies further showed that auditory cortical neurons processed the temporal modulation in a similar manner regardless of whether it was based on tone or noise carriers (Wang et al. 2002). Together, the evidence suggests that it is the "temporal modulation," and not the amplitude or frequency modulation per se, that most auditory cortical neurons appear to extract from a complex acoustic environment.

The similarity between cortical responses to sAM and sFM stimuli has been noted in some previous studies. Both the studies of Eggermont (1994) and that of Gaese and Ostwald (1995) found that there was no significant difference between $tBMF_{sAM}$ and $tBMF_{sFM}$. Similar mean tBMFs have been reported when nonsinusoidal envelopes were used to modulate tone carriers (Eggermont 1994; Schreiner and Urbas 1988). Our findings are consistent with these results in that the BMF of most cortical neurons is relatively insensitive to the spectral contents with which temporal modulations are created. However, the range of tBMF, over which the match between sAM and sFM responses was observed, was higher and more spread out in A1 under the awake condition as reported here than those observed in anesthetized animals in previous reports. Furthermore, we showed, on the basis of well-isolated single-units, that the similarity between cortical responses to sAM and sFM stimuli was also reflected in other response measures such as rBMF, BW, Q, and f_{max}.

Rate vs. temporal coding

While temporally modulated signals have been used in a number of previous studies of the auditory cortex in various species, some fundamental questions remain largely unanswered. An important issue is the role of rate coding in neural representation of temporally modulated signals in the auditory cortex. Previous studies have focused mainly on analysis of stimulus-synchronized temporal discharge patterns that are found throughout the auditory system. Under commonly used anesthetized conditions, there are very limited stimulus-driven sustained discharges in the auditory cortex to support rate-coding schemes. As the results of this study and other studies (e.g., Bieser and Müller-Preuss 1996) showed, the unanesthe-
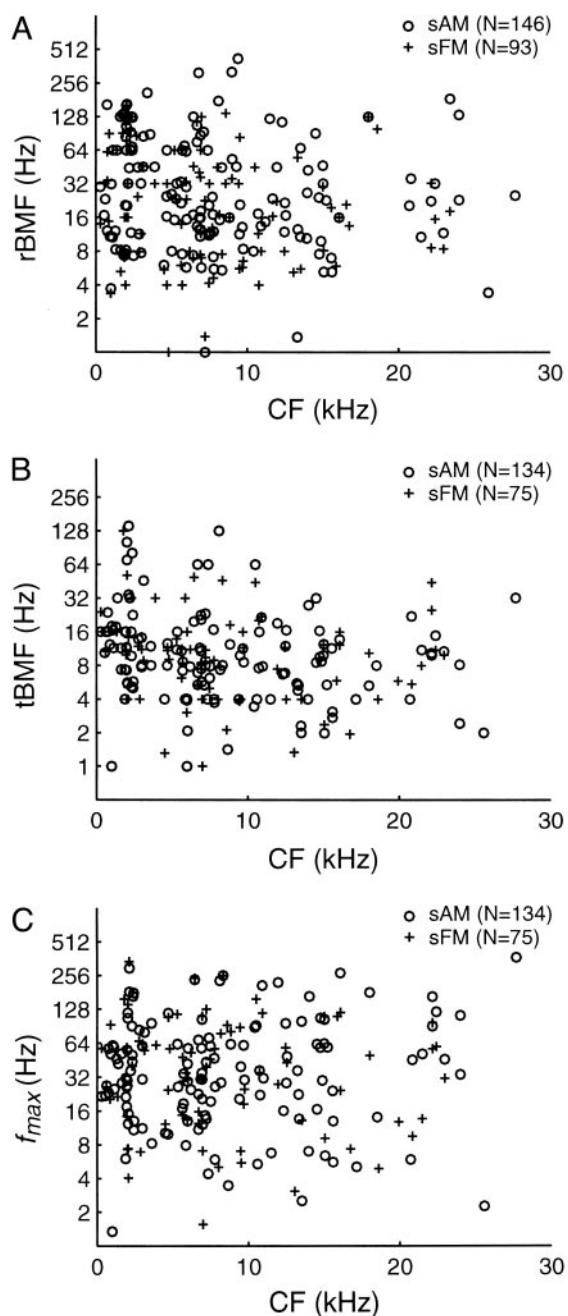
FIG. 20. Relationships between measures of MTF and CF of units are shown for both sAM (○) and sFM (+) data. *A*: rBMF, *B*: tBMF, and *C*: f_{max}. No significant differences were found between the medians calculated from units with CF less than the median CF and those with CF greater than the median CF for all 3 response measures (Table 3).

tized auditory cortex was highly responsive, and mean firing rate clearly demonstrated feature selectivity. The selectivity for a particular temporal modulation frequency in terms of average discharge rate in the absence of stimulus-synchronized discharge patterns was often observed in our recordings from the unanesthetized auditory cortex and represented a temporal-to-rate transformation by cortical neurons. The results from the present study suggest that the discharge rate-based temporal modulation selectivity may be involved in processing the finer temporal structures of complex sounds represented by higher temporal modulation frequencies that are not adequately en-

coded by stimulus-synchronized discharge patterns of cortical neurons. These findings provide further support for a two-stage temporal processing model we have proposed on the basis of cortical responses to sequential stimuli (Lu et al. 2001b). Bieser and Müller-Preuss (1996) suggested that high AM frequencies may be encoded in a rate code. We have further extended this notion by showing that the rate-coding is likely needed for encoding *temporal* modulations at high frequencies, in both the amplitude and frequency domain. In both the present study and that of Bieser and Müller-Preuss (1996), the carrier frequency of a sAM stimulus was set at a unit's CF. The notion of rate coding supported by these studies is conceptually different from the one suggested by Schulze and Langner (1997). In that study, it was shown that the firing rate of a cortical neuron could be influenced by high-frequency AM introduced at a carrier frequency far away from a unit's CF (Schulze and Langner 1997), a phenomenon that is generally interpreted as a spectral rather than temporal effect.

Modulation selectivity and onset timing

The findings in this study also bear some relationship to previous studies of onset response timing to acoustic stimuli as the ability of cortical neurons to represent periodic signals may be partially correlated with their minimum spike latency (e.g., Fig. 19). Although cortical neurons can show almost the same preciseness in their first spike latency, albeit with longer mean latencies, compared with the auditory-nerve (Heil 1997; Heil and Irvine 1997; Phillips and Hall 1990), they have a much diminished capacity to synchronize to rapidly occurring sequential acoustic events, such as the modulation cycles in the sAM and sFM stimuli. There is an important difference between the response to stimulus onset and that of a sustained periodic signal. In the latter case, the neuron may not have time to "reset" before the next acoustic event. Thus any long-lasting effects could contribute to the next response. It is possible that the peak in the rMTF or tMTF could be partially a result of a limited window of integration of responses to successive stimulation. Our data indicated that there were major differences between onset and sustained portions of the response (e.g., Figs. 2 and 17). We found that the bandwidth of rMTFs was generally narrower when considering mainly the sustained part of the response (e.g., Fig. 2*E*). The majority of the existing electrophysiological studies that investigated onset responses (including Heil and Irvine 1997; Phillips and Hall 1990) were typically conducted in anesthetized conditions where the responses were largely described as phasic. The onset responses characterized under such experimental conditions may not fully account for the variety of responses—especially rate-based, nonstimulus–synchronized responses—under unanesthetized conditions such as those reported here. For example, in the auditory cortex of awake animals, some cortical neurons do not exhibit well-timed onset responses (e.g., Fig. 2*C*).

Comparison with previous studies

Most of the previous single- or multi-unit studies of cortical representations of sinusoidal amplitude and frequency modulations were conducted using anesthetized animals with a few exceptions (Bieser 1998; Bieser and Müller-Preuss 1996; Creutzfeldt et al. 1980). Under anesthetized conditions, phasic

responses are the typical discharge patterns in the auditory cortex, even for responses to complex vocalizations. A distinct characteristic of neural responses in unanesthetized auditory cortex is sustained firing, a property that stands in sharp contrast to barbiturate-anesthetized cortex. It has long been recognized that anesthetics have profound side effects on the response properties of the auditory cortex, more so than other sensory cortices, possibly because of longer pathways leading to the cortex from the auditory periphery. Goldstein et al. (1959) showed that evoked cortical potentials could synchronize to higher repetition frequencies of clicks in the unanesthetized condition than in the anesthetized condition. Our recent studies showed that stimulus-following rates in A1 were higher in the awake marmosets than in anesthetized cats for click train stimuli (Lu and Wang 2000; Lu et al. 2001b).

The tBMFs observed in our study are higher than what has been observed in most studies of anesthetized animals and are comparable to other studies of A1 neurons in awake animals. Mean tBMFs obtained with sAM and sFM stimuli were 5.4 ± 3.4 and 5.9 ± 5.2 Hz, respectively, in the A1 of ketamine-anesthetized cats (Eggermont 1994) and ~ 10 Hz in the A1 of rats anesthetized by pentobarbital-based Equithesin (Gaese and Ostwald 1995). In both studies, tungsten microelectrodes were used and statistical significance of synchronized responses was assessed by Rayleigh test at $P < 0.001$, the same significance level used in the present study. A higher mean tBMF (14.2 ± 8.8 Hz) in A1 was reported by Schreiner and Urbas (1988) using sAM and rectangular AM stimuli. Their recordings, mostly multi-units, were made using low-impedance carbon fiber microelectrodes in paralyzed cats infused with pentobarbital anesthetics. The synchronization measure used in Schreiner and Urbas (1988) study included contributions of both the first and second harmonics of the modulation frequency from the Fourier spectrum. The spectral component corresponding to the first harmonic of the modulation frequency in their analysis is equivalent to the vector strength and similar measures used in the two studies cited above (Eggermont 1994; Gaese and Ostwald 1995). Bieser and Müller-Preuss (1996) studied cortical responses to sAM sounds in awake squirrel monkeys and reported a mean tBMF of 17.8 Hz in A1 with the distribution peaked between 8 and 16 Hz. The response synchronization in their study was quantified by cross-correlating the PSTH with the amplitude envelope of the AM stimuli. Multi-unit activities were recorded using low-impedance tungsten microelectrodes in this study. The Rayleigh test was not used to assess the statistical significance of response synchronization in the studies by Schreiner and Urbas (1988) and Bieser and Müller-Preuss (1996). The mean tBMF reported in the present study from awake marmosets is 15.6 ± 21.4 Hz for sAM and 14.2 ± 17.6 Hz for sFM stimuli, respectively (Table 1), calculated on a linear scale as has been done in the preceding referred studies.

In general, tBMF measured in awake animals has a greater spread (i.e., larger SD) than that measured in anesthetized animals, reflecting more diverse, stimulus-driven temporal discharge patterns. We did not observe in our data significant spontaneous oscillations as reported by other studies of anesthetized cortex (e.g., Eggermont 1992; Gaese and Ostwald 1995; Lu and Wang 2000). The BMFs derived from tMTF or rMTF reported in the present study clearly resulted from stimulus-driven activities. The lower values of tBMFs observed in

anesthetized auditory cortex (<10 Hz) suggested that they may possibly be correlated with cortical spindle frequencies under anesthesia (Kenmochi and Eggermont 1997). In fact, Gaese and Ostwald (1995) reported that tBMFs were significantly correlated with oscillation frequencies. Our recordings were based on well-isolated single-units (Fig. 1), obtained using high-impedance tungsten microelectrodes and were mostly from neurons in upper cortical layers that represent the output of A1. Most of the recordings in A1 of anesthetized animals were obtained using low-impedance electrode from middle cortical layers that receive the bulk of thalamocortical inputs. It is known that neurons in the medial geniculate body have much higher stimulus-following response rates (Creutzfeldt et al. 1980; de Ribaupierre et al. 1980).

We observed in the present study that rBMF was higher than tBMF in both individual neurons and across populations of cortical neurons (Fig. 11). This is consistent with observations made earlier by Schreiner and Urbas (1988) in cats and Bieser and Müller-Preuss (1996) in squirrel monkeys. Gaese and Ostwald (1995) found that most neurons in anesthetized rats did not display tuned rBMF. The distribution of rBMF (Fig. 5) obtained in the present study is similar to that reported by Bieser and Müller-Preuss (1996) (for sAM stimuli); both are centered at higher modulation frequencies and are more widely distributed across modulation frequency than those observed in anesthetized animals. There is also a significant difference between population-averaged rMTFs observed in A1 of anesthetized and awake animals. In the study of Eggermont (1994), the population-averaged rMTF has a peak near ~ 10 Hz, whereas our data show peaks between 16 and 32 Hz, about one octave higher (Fig. 12, *A* and *B*). This clearly indicates that the A1 is maximally excited at distinctly different modulation frequencies under the anesthetized and awake conditions.

Correlation with psychophysics

It has been shown that both humans (Viemeister 1979) and nonhuman primates (Moody 1994) are able to detect AM of a tone or noise carrier. However, like all other sensory systems, the auditory system has limited time resolution and is unable to detect amplitude changes that occur too rapidly. The temporal modulation transfer function is usually measured psychophysically based on the detection threshold, with respect to modulation depth, as a function of modulation frequency. The shape of a psychophysical tMTF is generally low-pass for low-frequency carriers up to ~ 60 Hz (for 1 kHz carriers) (Viemeister 1979), above which the tMTF begins to increase because sidebands in the frequency spectrum of the stimuli can then be resolved by the auditory system. As carrier frequency increases, the cutoff frequency of tMTF increases as well until $100–130$ Hz (for >2 kHz carrier frequency) (Kohlrausch et al. 2000) before spectral effects confound the measurement. Increased cutoff frequency of tMTF at higher carrier frequencies reflects increased width of the critical band (Zwicker and Fastl 1999). The majority of the neurons we studied had their CFs >1 kHz (Fig. 20), with the CF distribution centered near 7 kHz. The low-pass portion of psychophysical tMTF is considered a consequence of temporal processing, whereas the high-frequency portion is due to spectral effects. The cortical responses analyzed in the present study are likely correlated with the temporal aspect of the psychophysical tMTF. We did

observe neurons, in particular those in the non-BP group, whose discharges increased with increasing modulation frequency above ~100 Hz (Fig. 12, *A* and *B*).

The two measures of discharge synchrony used in this study are tBMF and f_{max}. The f_{max} measure should be more closely related to psychophysical tMTF because it reflects a threshold value of discharge synchrony rather than a maximum value as does the tBMF. The presence of discharge synchrony can be a basis for the auditory cortex to temporally discriminate modulated tones at low modulation frequencies versus unmodulated tones with the same carrier frequency. The cumulative distributions of f_{max} have a low-pass shape for both sAM and sFM stimuli (Fig. 9*B*) and begin to drop at ~4-Hz modulation frequency. There are <10% of neurons with significant discharge synchrony at modulation frequencies >128 Hz in our samples (Figs. 9*B*, 12 *C* and *D*). These characteristics mirror those of the psychophysical tMTF described in the preceding text. Furthermore, the psychophysical tMTF has been shown to be independent of sound level except at very low intensities. The tBMFs measured in cortical neurons are largely independent of sound level as well (Figs. 13 and 14).

In nearly all of the analyzed response measures, we found similarities between those obtained using sAM and sFM stimuli. This includes rBMF (Fig. 5), tBMF (Fig. 8), and f_{max} (Fig. 9) as well as BW and Q (Fig. 6). Such a similarity in cortical coding of these two classes of modulated sounds may underlie the similarity in the perception of these sounds (Edwards and Viemeister 1994; Moore and Sek 1992). It has been proposed by a number of psychophysical studies that FM signals are converted into an AM-equivalent representation in the auditory system. While both the auditory periphery and brain stem have been suggested as potential sites of such conversions on the basis of cochlear filters (Zwicker and Fastl 1999) and binaural processing (Saberi and Hafter 1995), respectively, it is unclear whether sAM and sFM stimuli like those used in the present study are similarly represented in these subcortical processing stations in the manner reported here. The findings of the present study show that the modulation frequencies embedded in sAM and sFM sounds are extracted and represented in a strikingly similar way by populations of auditory cortical neurons. They suggest that there might be a shared temporal processing mechanism at the level of both single and populations of cortical neurons.

Another possible psychophysical correlate of f_{max} is the lower limit of pitch, which is defined as the lowest repetition rate that evokes a sensation of pitch. Krumbholz et al. (2000) showed a lower limit of pitch near 30 Hz as did a companion study by Pressnitzer et al. (2001) using melody sequences. The distributions of f_{max} are centered close to 30 Hz for the cortical neurons we studied (Fig. 9, Table 1). For modulation frequencies below f_{max}, modulation periods can be resolved temporally because there are significant discharges synchronized to the modulation periods. Higher modulation frequencies that correspond psychophysically to the sensation of roughness (Zwicker and Fastl 1999), for example, are likely represented by a rate code instead of by a temporal code because the rBMF can be as high as 256 Hz in individual neurons (Fig. 5). Our recent study using click train stimuli suggested rate-coding for even higher repetition frequencies (Lu et al. 2001b).

Finally, it would be interesting to know whether the preferred temporal modulation frequencies in the A1 of awake marmoset monkeys are comparable to those of other primate species including humans under unanesthetized conditions.

REFERENCES

BIESER A. Processing of twitter-call fundamental frequencies in insula and auditory cortex of squirrel monkeys. *Exp Brain Res* 122: 139–148, 1998.

BIESER A AND MÜLLER-PREUSS P. Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 108: 273–284, 1996.

BLACKBURN CC AND SACHS MB. Classification of unit types in the anteroventral cochlear nucleus: post-stimulus time histograms and regularity analysis. *J Neurophysiol* 62: 1303–1329, 1989.

CREUTZFELDT O, HELLWEG F-C, AND SCHREINER C. Thalamocortical transformation of responses to complex auditory stimuli. *Exp Brain Res* 39: 87–104, 1980.

DE RIBAUPIERRE F, ROUILLER E, TOROS A, AND DE RIBAUPIERRE Y. Transmission delay of phase-locked cells in the medial geniculate body. *Hear Res* 3: 65–77, 1980.

EDWARDS BW AND VIEMEISTER NF. Modulation detection and discrimination with three-component signals. *J Acoust Soc Am* 95: 2202–2212, 1994.

EGGERMONT JJ. Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hear Res* 56: 153–167, 1991.

EGGERMONT JJ. Stimulus induced and spontaneous rhythmic firing of single units in cat primary auditory cortex. *Hear Res* 61: 1–11, 1992.

EGGERMONT JJ. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform, and intensity. *Hear Res* 74: 51–66, 1994.

EVANS EF AND WHITFIELD IC. Classification of unit responses in the auditory cortex of the unanesthetized and unrestrained cat. *J Physiol (Lond)* 171: 476–493, 1964.

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.

GAESE BH AND OSTWALD J. Temporal coding of amplitude and frequency modulation in the rat auditory cortex. *Eur J Neurosci* 7: 438–450, 1995.

GAESE BH AND OSTWALD J. Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. *J Neurophysiol* 86: 1062–1066, 2001.

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.

GOLDSTEIN MH, JR, KIANG NY-S, AND BROWN RM. Responses of the auditory cortex to repetitive acoustic stimuli. *J Acoust Soc Am* 31: 356–364, 1959.

HEIL P. Auditory cortical onset responses revisited. I. First-spike timing. *J Neurophysiol* 77: 2616–2641, 1997.

HEIL P AND IRVINE DR. First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78: 2438–2454, 1997.

HOUTGAST T AND STEENEKEN HJM. The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acustica* 28: 66–73, 1973.

JOHNSON DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* 68: 1115–1122, 1980.

JORIS PX AND YIN TCT. Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* 91: 215–232, 1992.

KENMOCHI M AND EGGERMONT JJ. Autonomous cortical rhythms affect temporal modulation transfer functions. *Neuroreport* 8: 1589–1593, 1997.

KOHLRAUSCH A, FASSEL R, AND DAU T. The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am* 108: 723–734, 2000.

KRUMBHOLZ K, PATTERSON RD, AND PRESSNITZER D. The lower limit of pitch as determined by rate discrimination. *J Acoust Soc Am* 108: 1170–1180, 2000.

LANGNER G. Periodicity coding in the auditory system. *Hear Res* 60: 115–142, 1992.

LANGNER G AND SCHREINER CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 60: 1799–1822, 1988.

LIANG L, LU T, SNIDER RK, AND WANG X. Cortical response to amplitude-, frequency-, and mixed-modulation in the awake marmoset monkeys. *Assoc Res Otolaryngol Abstr* 22: 35, 1999.

LU T, LIANG L, AND WANG X. Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *J Neurophysiol* 85: 2364–2380, 2001a.

LU T, LIANG L, AND WANG X. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4: 1131–1138, 2001b.

LU T AND WANG X. Temporal discharge patterns evoked by rapid sequences of wide- and narrowband clicks in the primary auditory cortex of cat. *J Neurophysiol* 84: 236–246, 2000.

MARDIA KV AND JUPP PE. *Directional Statistics*. New York: Wiley, 2000.

MOODY DB. Detection and discrimination of amplitude-modulated signals by macaque monkeys. *J Acoust Soc Am* 95: 3499–3510, 1994.

MOORE BC AND SEK A. Detection of combined frequency and amplitude modulation. *J Acoust Soc Am* 92: 3119–3131, 1992.

PALMER AR. Encoding of rapid amplitude fluctuations by cochlear-nerve fibers in the guinea pig. *Arch Otorhinolaryngol* 236: 197–202, 1982.

PFINGST BE AND O'CONNOR TA. Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J Neurophysiol* 45: 16–34, 1981.

PHILLIPS DP AND HALL SE. Response timing constraints on the cortical representation of sound time structure. *J Acoust Soc Am* 88: 1403–1411, 1990.

PRESSNITZER D, PATTERSON RD, AND KRUMBHOLZ K. The lower limit of melodic pitch. *J Acoust Soc Am* 109: 2074–2084, 2001.

RALL W AND AGMON-SNIR H. Cable theory for dendritic neurons. In: *Methods in Neuronal Modeling*, edited by C. Koch and I. Segev. Cambridge, MA: The MIT Press, 1998, p 27–92.

RECANZONE GH, GUARD DC, AND PHAN ML. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 83: 2315–2331, 2000.

DERIBAUPIERRE F, GOLDSTEIN MH JR, AND YENI-KOMISHIAN GH. Cortical coding of repetitive acoustic pulses. *Brain Res* 48: 205–225, 1972.

RICE JA. *Mathematical Statistics and Data Analysis*. Pacific Grove, CA: Wadsworth and Brooks, 1988.

ROSEN S. Temporal information in speech: acoustic, auditory, and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336: 367–373, 1992.

SABERI K AND HAFTER ER. A common neural code for frequency- and amplitude-modulated sounds. *Nature* 374: 537–539, 1995.

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.

VIEMEISTER NF. Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am* 66: 1364–1380, 1979.

WANG X. On cortical coding of vocal communication sounds in primates. *Proc Natl Acad Sci USA* 97: 11843–11849, 2000.

WANG X, LIANG L, ELIADES S, SNIDER RK, LU T, KADIA SC, AND BARBOUR D. Characteristics of single-unit responses in the auditory cortex of awake marmoset monkeys. *Assoc Res Otolaryngol Abstr* 22: 35, 1999.

WANG X, LU T, AND LIANG L. Temporal processing in the primate auditory cortex. In: *Physiological and Psychophysical Bases of Auditory Function. Proceedings of the 12th International Symposium on Hearing*, edited by Houtsma AJM, Prijs VF, and Schoonhoven R. The Netherlands: Shaker Publishing BV, 2001, p. 352–357.

WANG X, LU T, AND LIANG L. Cortical processing of temporal modulations. *Speech Commun* In press.

WANG X AND SACHS MB. Transformation of temporal discharge patterns in a VCN stellate cell model: implications for physiological mechanisms. *J Neurophysiol* 73: 1600–1616, 1995.

WHITFIELD IC AND EVANS EF. Responses of auditory cortical neurons to stimuli of changing frequency. *J Neurophysiol* 28: 655–672, 1965.

ZWICKER E AND FASTL H. *Psychoacoustics*. Berlin: Spinger, 1999.